Title
Unions, Job Training, and the Wages of Foreign-Born Workers in the U.S.

Author
Chen, Hung-Lin

Publication Date
2010

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Unions, Job Training, and the Wages of Foreign-Born Workers in the U.S.

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Economics

by

Hung-Lin Chen

December 2010

Thesis Committee:
Dr. David Fairris, Co-Chairperson
Dr. Todd Sorensen, Co-Chairperson
Dr. Mindy Marks

The Dissertation of Hung-Lin Chen is approved:

Committee Co-Chairperson

Committee Co-Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

I would like to thank my two advisors, Dr. David Fairris and Dr. Todd Sorensen, for their support, encouragement, guidance, and patience. They always spent time reading my papers carefully and made time for me to discuss my dissertation in their busy schedules. For my future career, they gave important suggestions and help. When I felt helpless, they were always there to solve my predicament. Without their brilliant guidance, this dissertation could not be done. Without their encouragement and support, I could not survive in our graduate program. I truly consider myself extremely lucky to have them as my dissertation advisors.

I am also deeply grateful to my other committee members, Dr. Aman Ullah and Dr. Mindy Marks, for their invaluable advice regarding my research. Besides research, Dr. Mindy Marks greatly helped me with preparing job market and Dr. Jang-Ting Guo constantly helped me during these graduate years. I am truly indebted to all of them.

I would like to give a special thank to my parents, Ming-Fu Chen and Chin-Man Lin. Without their unconditional support, this dissertation could not be completed. I also would like to thank my boyfriend, Kohei Takahashi, for always being there for me with support and encouragement during stressful times.

For those friends, I am thankful to Trinidad Beleche, Sarah Daway, Daniel Farhat, Insu Kim, Jie Li, Xiangbo Liu, Dackie Wang, James Wu, Yuan Xu, and all others at University of California, Riverside. My sincere gratitude is expressed to all of them.

Finally, I thank all the staff members at the department of economics, who have been very friendly and helpful.

DEDICATION

To my parents

ABSTRACT OF THE DISSERTATION

Unions, Job Training, and the Wages of Foreign-Born Workers in the U.S.

by

Hung-Lin Chen

Doctor of Philosophy, Graduate Program in Economics
University of California, Riverside, December 2010
Professor David Fairris and Professor Todd Sorensen, Co-Chairpersons

According to the U.S. Bureau of Labor Statistics, the share of foreign-born workers in the labor market increased from 14% to 17% between 1994 to 2008. At the same time, foreign-born union members increased, from 9% to 11%. Immigrants in the United States are an economically disadvantaged group in the labor market. Previous studies suggest that union members and other workers covered by collective agreements receive union wage premiums about 15 percent over nonunion members in the United States. Joining unions could be a good approach for foreign-born workers to receive higher wages. In this connection, the goal of this paper is twofold: one, to estimate the willingness of foreign-born workers to join unions, and two, to determine the union wage premium for foreign-born workers and whether there is a statistical difference in the union relative wage effect for foreign- and native-born workers. The results show that foreign-born workers have a lower probability of joining unions, ceteris paribus. The wage differential between union and nonunion workers for foreign-born workers is only 11.3%, while that for native-born workers is 13.3%. This 2-percent difference of the union impact on wages of native- and foreign-born workers is statistically significant. Among the foreign-born workers, Mexicans have the highest union relative wage effect (22.4%). This study also finds that the union/nonunion wage differentials for both female foreign- and native-born workers are smaller than those for their

male counterparts. Moreover, the union wage premium is greater for foreign-born workers in the private sector than for those in the public sector. By region, unions have higher wage impact in the West Coast than in the East Coast.

In light of the numerous criticisms leveled against estimating the wage differential between union and nonunion workers using the ordinary least squares (OLS) method, this study estimates the union impact on wages of foreign-born and native-born workers using the propensity score matching (PSM) methodologies (nearest neighbor and kernel matching methods) proposed by Rosenbaum and Rubin (1983), and compares the results with those obtained using the OLS approach. The data on the wages and salaries of male workers aged 16 years and above are obtained from the Current Population Survey and span the period 1994-2008. Both the propensity score matching and OLS estimates indicate that the union/nonunion wage differentials for male foreign-born workers lie between 12% (OLS) and 27% (PSM). In addition, our results suggest that there is little difference in the union/nonunion wage differential between native- and foreign-born workers. The estimates of the union impact on wages using the propensity score matching technique are higher than those derived using OLS for both native- and foreign-born workers. Furthermore, among the foreign-born workers, the union relative wage effect is found to be higher for Mexican-born workers (26-42%), while for Asian-born workers it is lower and statistically insignificant based on the OLS. Further decomposition of the data into three different skill groups (high school dropouts, those with a high school degree, and those with a college degree or higher) reveals that, in general, the less skilled (high school dropout) foreign-born workers have the greatest union wage impact. However, the union wage premium is relatively larger for highly skilled workers (with a college degree or higher) among Mexican-born workers.

In chapter 3, we study the effect of job training on the US immigrant workers, using the 1996, 2001 and 2004, Survey of Income and Program Participation (SIPP) data. We improve upon prior studies by setting up our training evaluation model, studying the impact of training on both the average and the distributional earning of workers, and comparing the differences in the return to training for immigrant and native workers by applying the

Quantile regression (QREQ) model, the DiNardo, Fortin and Lemieux (DFL) reweighting methods, and propensity score matching method (PSM). From our distribution study, we find that training has a positive effect on wages for immigrant workers for most parts of income distribution. The DFL reweighting technique shows that after removing all observable characteristics differences between trained and untrained workers, training still increases wage premium for both natives and immigrants throughout the income distribution. Our analysis provides strong evidence for the hypothesis that after corrected for observable characteristics differences between trained and untrained workers, the effect of training is relatively larger for rich natives, much larger for middle income natives and similar for the poor natives and immigrants. Furthermore, the PSM results show that the job training premium for foreign-born workers is between 0.063 and 0.184, whereas for native-born workers it is between 0.108 and 0.229. There is 4-percent difference in the job training premium between native and immigrant. All estimates are statistically significant. Our results suggest that OLS estimates underestimate the training premium.

Contents

List of Tables

List of Figures

Chapter 1

Unions and Foreign-Born Workers in the U.S.

1.1 Introduction

There have been a growing number of foreign-born workers in the United States over the past few decades. According to the U.S. Bureau of Labor Statistics, the share of foreign-born workers in the labor market increased from 14% to 17% between 1994 to 2008. At the same time, foreign-born union members increased, from 9% to 11%. Immigrants in the United States are an economically disadvantaged group. On average, their wages are lower. Previous studies suggest that union members and other workers covered by collective agreements receive union wage premiums about 15 percent over nonunion members in the United States. Joining unions could be a good approach for foreign-born workers to receive higher wages. Thus, it is important for labor researchers to know how foreign-born workers are affected by unions and what size of union wage premium is for foreign-born workers.

The empirical literature on the union relative wage effects in the United States is extensive. Some of the studies concern about the union relative wage effects across gender and ethnicity. For instance, in his influential survey of the U.S. studies, Lewis (1986) concludes that there is very little difference between the union wage premium for male and female workers and it is not clear that whether there is a difference between the union wage premium for white and nonwhite workers. However, there is no research focusing on the effects of unions on the immigrant wage to date. Since there is an increasing share of immigrant workers in the labor market and an increasing share of union members who are foreign-born workers, the roles of immigrants in the labor market are getting important. As labor economists, we have to know that whether unions help foreign-born workers to raise wages in the United States. This paper is the first study on the union wage impact for foreign-born workers and also the first study to compare the union wage premium for foreign-born workers with that for native-born workers. We also explore the union wage premiums for those foreign-born workers who are from different birthplaces, who reside at different regions, and who work in private and public sectors.

Policy makers might be interested in immigrants since this group is getting larger in the United States. Joining unions is one of the approaches to increase immigrants' earnings. A better earning capability of immigrants would likely result in a reduction of social problems such as unemployment, gangs and crimes activities. Nonexperimental data have shown that increased earnings reduce criminal activity (Lalonde, 1986) and lower the number of crimes such as murders (Donohue and Siegelman, 1998). In addition, AFL-CIO doesn't like temporary worker programs and foreign-born workers to protect U.S. labor markets (Roy Beck, President, NumbersUSA.com). Oppositely, some unions supported temporary worker programs, such as Service Employees International Union (SEIU) and UNITE HERE (Fan and Batalova, 2007). Hence, to help policy makers improve their decisions for immigrant workers, unions, and the society, we need to gain better understanding of the relationship between unions and foreign-born workers.

In this paper, firstly, I examine the willingness of foreign-born workers to join unions. The results show that foreign-born workers have a lower probability of joining unions, ceteris paribus. The wage differential between union and nonunion workers for foreign-born workers is 12%, while that for native-born workers is 14%. This 2-percent difference of the union impact on wages of native- and foreign-born workers is statistically significant. That is, the Chow test rejects the null hypothesis of equality of the coefficients for the both worker groups at the 0.1% level of significance. Among foreign-born workers, Mexicans have the highest union relative wage effect (25.1%). This paper also finds that the union/nonunion wage differentials for both female foreign- and native-born workers are smaller than those for their male counterparts. Moreover, the union wage premium is greater for foreign-born workers in the private sector than for those in the public sector.

The rest of the paper proceeds as follow: section 1.2 reviews the relevant literature, section 1.4 describes the data, section 1.3 describes my estimation equations and identification strategy, section 1.5 presents the results, and section 1.6 concludes.

1.2 Literature Review

The estimates of union wage impact have been found to be positive and vary by countries with different income level. In the forth chapter of their book, Aidt and Tzannatos (2002) present summary estimates of the union wage premium for six high-income countries, four middle-income countries, and one low-income country. They suggest that higher-income countries have lower union relative wage effects than middle- and low-income countries. In the United States (high-income country), there have been more than 200 studies on union wage impact and the average estimate of union wage premium is approximately 15 percent (Lewis, 1986; Blanchflower and Freeman, 1996; Blanchflower, 1997). There are also many studies on this topic in the United Kingdom and the estimated union wage premium is approximately 10 percent (Booth, 1995; Blanchflower, 1997). The studies for middle- and low-income countries are limited. For example, Teal (1996) uses three-year surveys of manufacturing firms in Ghana (low-income country) and suggests that the union wage premium varies between 21 and 28 percent. These estimates are higher than those in high-income countries.

There has been a substantial literature on union/nonunion wage differentials across different types of workers. Lewis (1986) surveys 39 studies about union/nonunion wage differentials between female and male workers and concludes that there are very little union relative wage differentials for female and male workers in the United States. However, some studies show that the impact of unions on women's wages is greater than that on men's wages in Britain, Japan, South Africa, and Mexico (Main and Reilly, 1992; Standing, 1992; Moll, 1993; Panagides and Patrinos, 1994; Blanchflower and Freeman, 1996; Blanchflower, 1997; k. Nakumura et al., 1988; Aidt and Tzannatos, 2002). In Japan, k. Nakumura et al. (1988) suggest that the union/nonunion wage differentials is 10 percent for women but find nothing for man. In Mexican studies, Panagides and Patrinos (1994) estimate that the union/nonunion wage differentials for female workers is 9.8 percent points higher than that for male workers.

Some studies estimate the gender gap in the union and nonunion sectors instead of estimating union/nonunion wage differentials for female and male workers. Simpson (1985) finds that the gender gap is 22.9 percent in the union sector and 20.3 percent in the nonunion sector in Canada. His conclusion is that unions have little impact on the gender wage gap. However, Fairris (2003) uses Mexican's data and suggests that it is statistically insignificant gender discrimination in pay in the union sector in the 1984 results. In the nonunion sector, the gender wage gap is larger and statistically significant.

For race, Lewis (1986) surveys 46 studies in the United States and concludes that it is not clear whether there is a substantial difference between the wage markup for white and nonwhite workers; even so, Blanchflower (1997) suggests that nonwhite unionized workers have higher union wage premiums than white unionized workers in the United Kingdom. Dabalen (1998) finds that white workers (about 30 percent) have higher union relative wage effects than black workers (about 16 to 20 percent) in South Africa. In Mexico and Canada, unions have been found to reduce the discrimination against indigenous workers (Patrinos and Sakellariou, 1992; Panagides and Patrinos, 1994)

Previous union studies also have done some estimates of union/nonunion wage differentials for workers with different skills. They suggest that the union wage premium is usually higher for less skilled workers in Canada, United Kingdom, and United States (Lewis, 1986; Freeman and Medoff, 1984; Simpson, 1985). However, there is a different result in the study of South Africa. Unskilled nonwhite workers have the highest union wage premium, but semiskilled and skilled white workers have higher union wage premium (Moll, 1993). Dabalen (1998) estimates union wage premium in South Africa five years later and confirms the results suggested by Moll (1993).

The union relative wage effects are higher in the private sector than in the public sector in the United States, United Kingdom, and Canada (Aidt and Tzannatos, 2002). In the United Kingdom, Green (1988) finds that the wage markup is smaller in the public sector than in the private sector . However, Blanchflower (1997) concludes that the estimates are very similar in the private and public sectors in the United Kingdom. Simpson (1985) suggests a similar results in Canada.

Briggs (2001) mentions in his book that post-1965 immigrants have less and less education level, and the wage differential between foreign-born workers and native-born workers has become larger and larger. The average earning of men from Mexico and Central America were about half of those of native men. Even though they had similarly working experience as native men, they would have earned 27 percent less than native men.

Capps et al. (2003) have some key findings for immigrants in their paper. Most of the foreign-born workers dominate in the low-wage jobs especially male immigrants and their wages are even much lower than the minimum wage. These foreign-born workers with low wages don't speak English well. Two of every five low-wage foreign-born workers are undocumented. Most of the foreign-born workers are from Mexico (28%), and Asia (23%). Congressional Budget Office also finds that six states, California, Texas, Florida, New York, New Jersey and Illinois are foreign-born workers' favorite. Mexican-born workers are more likely to work in production, construction and extraction and building and grounds cleaning and maintenance and other foreign-born workers have broadly distributed occupations.

Although the literature on the union/nonunion wage differentials is substantial, those studies only focus on gender and ethnic group. Since there is an increasing share of immigrant workers in the labor market and an increasing share of union members who are foreign-born workers, it is important for researchers to know that do unions really help foreign-born workers.

1.3 Empirical Model

In this section I describe the empirical techniques which are used to answer the research questions laid out in the introduction. I begin by modeling the probability of foreign-born workers joining unions, and then estimate the difference in the native-immigrant wage differentials in the union and nonunion sectors. Finally, I estimate the difference in the union wage differentials between native-born and foreign-born workers.

1.3.1 Union Membership as a Choice Variable

Here, I estimate the likelihood of joining unions for foreign-born workers. Equation 1.1 gives a model of an underlying latent variable for being in the union sector.

$$U_i^* = \phi_0 + \phi_M M_i + \phi_X X_i + \nu_i \tag{1.1}$$

where U_i^* is the latent variable of union coverage for individual i, M is an indicator variable equal to 1 if this person is a foreign-born workers and 0 otherwise, X is a vector of other observed union membership determining characteristics, the ϕ terms represent the associated coefficients, and ν_i is the unobserved characteristics that determine union membership. The sign of ϕ_M shows the relative probability of union membership for foreign-born workers and is the coefficient of interest.

The relationship between union coverage and the latent variable is as follows:

$$U = \begin{cases} 1 & \text{if} \quad U^* > 0 \quad \text{Union worker} \\ 0 & \text{if} \quad U^* \leq 0 \quad \text{Nonunion worker} \end{cases}$$

1.3.2 The Native-Immigrant Wage Gap: Differences by Union Sector

Here I attempt to determine whether foreign-born workers fare differently, when compared to native-born workers, in the union sector than they do in the nonunion sector. Since there is a different wage determination process for unionized and nonunionized workers, more recent estimates are based on the two separate regressions(Aidt and Tzannatos, 2002). Farber (2001) also mentioned that "unions tend to standardize wages and attach wages to jobs, it is likely that the function characterizing the earning of union workers will be flatter in skill dimensions than the function for nonunion workers". Therefore, I model two log wage equations for union and nonunion sectors separately below, with W again representing log wages, M an indicator variable for immigrants, and X a vector of other characteristics determining wage outcomes. The coefficients on these characteristics are now given by the δ terms, and unobserved characteristics are represented by υ.

$$W_{i_{(NU)}} = \delta_{0_{(NU)}} + \delta_{M_{(NU)}} M_{i_{(NU)}} + \delta_{x_{(NU)}} X_{i_{(NU)}} + \upsilon_{i_{(NU)}} \quad (1.2)$$

$$W_{i_{(U)}} = \delta_{0_{(U)}} + \delta_{M_{(U)}} M_{i_{(U)}} + \delta_{x_{(U)}} X_{i_{(U)}} + \upsilon_{i_{(U)}} \quad (1.3)$$

In equation 1.2, I estimate a wage equation for the nonunion sector,NU, and in equation 1.3, I do the same for the union sector, U. The coefficients of $\delta_{M_{(NU)}}$ and $\delta_{M_{(U)}}$ show the native-immigrant wage gap in the nonunion and union sectors and they are the coefficients of interest.

1.3.3 The Union Wage Premium: Difference by Nativity

In this section, I develop a model to estimate the union wage premiums for native- and foreign-born workers. According to the previous literature, there are two approaches to estimate union impact on wages for different types of workers. For example, One approach is to pool all observations in the union and nonunion sectors and run a regression of log wages on other characteristics determining wage outcomes plus an dummy variable indicating union times an indicator variable for immigrants. The second approach is to use separate regressions for union and nonunion workers and see the coefficients of immigrant variable. Many researchers argue that there is a problem with the former approach. It assumes that the marginal effect of skill on earnings is the same in both groups of workers and it is not true in practice.

Since the marginal effect of skill on earnings for native- and foreign-born workers is different, we generate two equations for wages that vary across two groups, where the groups here are foreign-born workers (M) and native-born workers (N) and the key variable of interest is U, the indicator variable for union coverage.

$$W_{i_{(N)}} = \delta_{0_{(N)}} + \delta_{U_{(N)}} U_{i_{(N)}} + \delta_{x_{(N)}} X_{i_{(N)}} + e_{i_{(N)}} \tag{1.4}$$

$$W_{i_{(M)}} = \delta_{0_{(M)}} + \delta_{U_{(M)}} U_{i_{(M)}} + \delta_{x_{(M)}} X_{i_{(M)}} + e_{i_{(M)}} \tag{1.5}$$

Because the union variable is related to other observable and unobservable variables, equation 1.2 to equation 1.5 are likely to suffer from selection bias (Farber, 2001; Card, 1996; Hirsch and Schumacher, 1998; Budd and Na, 2000; Hildreth, 2000)[1] Lewis (1986) notes that it is unable to determine the size and direction of the selectivity bias.

[1]For instance, workers who are less productive (an unobserved variable) are more likely to join the unionized sector and will have a higher union relative wage effect compared to those who are more highly productive. In order to get consistent estimates of the relative union wage premium for both native- and foreign-born workers in this study, we assume the selectivity problem is the same for these two groups.

1.4 Data

The data for this analysis are obtained from the IPUMS-CPS (Current Population Survey), a monthly U.S. household survey conducted jointly by the U.S. Census Bureau and the Bureau of Labor Statistics from 1962 to 2008. It offers a large sample size and complete information on union membership and coverage since 1983, and has also included data identifying foreign-born workers since 1994. Therefore, our data are drawn from 1994 to 2008.

The foreign-born population includes immigrants, legal non-immigrants (e.g., refugees and persons on student or work visas) and persons illegally residing in the United States. The term "native-born worker" refers to people residing in the United States who are citizens of the United States in one of three categories: (1) people born in one of the 50 states and the District of Columbia; (2) people born in the United States insular areas, such as Puerto Rico or Guam; or (3) people who were born abroad to at least one parent who was a United States citizen.

All data refer to wage and salary workers who are employed and are 16 years of age and older. The dependent variable is log hourly wages and the variable of interest is union membership. The sample consists of 107,899 native-born workers and 15,751 foreign-born workers. It also consists of 102,619 nonunion workers and 21,031 union workers. Our analysis controls for the following covariates: potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 state effects, 15 year effects, 144 industry effects, and 229 occupation effects.

Table 1.1 presents the sample means for union and nonunion members. Mean hourly wage for nonunion workers amounts to 11.407 and for union workers amounts to 16.074. Union workers have higher unconditional union wage premiums about 41 percent. On average, union members have 4 more years of potential experience (age-education-6), less education, and more pension plan at work. They are more likely to be elder (4 years older), male, non-white, married, and native-born. Moreover, workers who residing in the northeast, midwest, west, metropolitan areas and working in the large firms and private sectors are more likely to join the union sectors.

Table 1.2 presents the sample means for native-born and foreign-born workers. Mean hourly wages amount to 10.923 and 9.230 for native-born and foreign-born workers, respectively. Native-born workers have higher unconditional wage about 18 percent. On average, foreign-born workers have 1 more year of potential experience (age-education-6), less pension plan at work, fewer years of completed education, and are more likely to be younger, married, and male. Moreover, foreign-born workers are more likely to reside in the northeast, west, and metropolitan areas and work in the relative small firms and private sectors.

1.5 Results

Table 1.3 presents the probability of joining unions for foreign-born workers. Foreign-born workers have less probability to join unions (the coefficient is -0.099 and statistically significant), especially for male foreign-born workers (the coefficient is -0.118 and statistically significant). Also, foreign-born workers who reside in the Midwest (-0.210) and West regions (-0.125), and work in the public sector (-0.155) have relatively less probability to join in the union sectors. This is consistent with the point that Waldinger and Der-Martirosian (2000) mentioned in their paper that employers are less likely to hire foreign-born workers and some of the reasons might be English and visa problems. Milkman (2007) also says that most of the foreign-born workers temporarily stay in the United States and they are not interested in joining unions and pay the union dues. To be union members, they have to pay full union dues and it costs 2 hours of wages per month (Budd and Na, 2000). However, Milkman (2007) suggested that Latinos are more positive toward unionism than whites. Some labor unions don't like temporary foreign-born workers. For instance, AFL-CIO doesn't like temporary worker programs and foreign-born workers to protect U.S. labor markets (Roy

Beck, President, NumbersUSA.com). Oppositely, some unions supported temporary worker programs, such as Service Employees International Union (SEIU) and UNITE HERE(Fan and Batalova, 2007).

Then, I examine the native-immigrant wage gaps in the union and nonunion sectors and union wage impact for native-born and foreign-born workers using four separate ordinary least squares (OLS) equations. Table 1.4 presents that foreign-born workers who worked in the nonunion sectors have 5.1% lower wages than native-born workers. In the union sector, the native-immigrant gap is slightly larger, 5.9%. Both coefficients are statistically significant but a Chow test fails to reject the null hypothesis of equality of the two coefficients. These results are similar to previous gender studies (Simpson, 1985; Farber, 2001). Simpson (1985) estimates the gender wage gap in the union and nonunion sectors in Canada. He found that the gender gap is 22.9 percent in the unionized sector and 20.3 percent in the nonunionized sector. His conclusion is that unions have little impact on the gender wage gap. Our results could also suggest that unions have little impact on the native-immigrant wage gap. In addition, Farber (2001) found that the sex differential in wages is larger in the union sector and concluded that females are substantially segregated in particular occupations, both within and outside the union sector. Our results are similar to Farber (2001)'s results. That could also be the reason that foreign-born workers are segregated in particular occupations like production, construction and extraction and building and grounds cleaning and maintenance.

The union/nonunion wage differential for native-born workers are 13.3% and the one for foreign-born workers are 11.3%. Native-born workers have 2% point higher union wage premiums than foreign-born workers and the difference is statistically significant. A Chow test rejects the null hypothesis of equality of the coefficients. All four regressions include potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 state effects, 15 year effects, 144 industry effects, and 229 occupation effects. Our results also show that all coefficients of human capital characteristics have the expected sign. Relative to native-born workers, foreign-born workers who have more than bachelor's degree have lower economic returns to years of education and experience.

I also estimate the native-immigrant wage gaps and the union/nonunion wage differentials for gender. The native-immigrant wage gaps for female workers in both union and nonunion sectors are relatively smaller than those for male workers (table 1.5). According to previous study, female foreign-born workers are better educated and more likely to reside in the United States legally (Capps et al., 2003). This could be the reasons that the native-immigrant wage gap for female workers is smaller. The native-immigrant wage gap for female workers in the nonunion sector is 3.3%, while that in the union sector is 5.6%. Again, our results show that unions have little impact on female native-immigrant wage gap. The union wage premiums are larger for male workers than that for female workers. The difference of union wage premium between female foreign-born and native-born workers is slightly smaller (1.8%) than for that between male foreign-born and native-born workers (2.7%).

Foreign-born workers who come from different countries have different union wage impact. Most of the foreign-born workers come from Central American, Asian, and Europe. About one half of foreign-born workers are Hispanics. Among foreign-born workers, Mexican-born workers are more likely to join the unions. Those who don't join unions have 11.9% lower wages than native-born workers (see table 1.6). If Mexican-born workers join unions, the wage difference between them and native-born workers becomes smaller (-5.9%). Our results show that unions have larger impact on native-mexican wage gap. The union/nonunion wage differential for Mexican-born workers is 22.4%, much higher than that for native-born workers (13.3%). However, the rest foreign-born workers who are from other birthplaces have opposite stories. Their union wage premiums are lower (Asian and European) or almost equal (those foreign-born workers from South America) than native-born workers. The union wage premium for those workers from Africa is insignificant. See table 1.6 and table 1.7.

In table 1.2, our results show that foreign-born workers are more likely to reside in the private sector than in the public sector. The conditional wage differences between foreign-born and native-born workers in both union and nonunion sectors in public sectors are relatively smaller than those in the private sectors (table 1.8). Our results also show that the union wage premium is 4-percent higher for both native- and foreign-born workers in the private sectors.

Table 1.9 to Table 1.11 present the union wage premiums in different regions. Foreign-born workers have higher union wage premiums in the west region (14.3%). Also, the difference of union wage premium between foreign-born and native-born workers is smaller in the west region (0.3%). In addition, those immigrants who work in the east and south regions have much lower union wage premiums. For those workers who are from Mexico working in the east, midwest and west regions, they have much higher union wage premiums than native-born workers do. Mexican-born workers working in the south region don't have significant union wage premium (-0.035). Asian-born workers don't have significant union wage premiums when they work in the east and midwest regions. The union wage premium for Asian is lower than that for native-born workers.

1.6 Conclusions

The results of this study suggest that foreign-born workers have a lower probability of joining unions, ceteris paribus. Native-immigrant wage gap is slightly larger in the union sector than in the nonunion sector. Our results suggest that unions have little impact on the native-immigrant wage gap. Moreover, foreign-born workers are segregated in some occupations.

The wage differential between union and nonunion workers for native-born workers is 13%, while that for foreign-born workers is only 11%. There is little difference between the union wage premium for native- and foreign-born worker. Our results are consistent with previous gender studies in the United States. Among the foreign-born workers, Mexicans have the highest union relative wage effect (22.4%). Unions have larger impact on the native-mexican wage gap.

Policy makers and researchers have been paying more attention on legal and illegal immigrants since this group has been getting larger and larger in the United States. Our results show that immigrants do gain more earnings when they are in the union or being paid by collective agreement. Unions do help immigrants to earn more. Previous literature show that a better earning capability of immigrants would likely result in a reduction of social problems such as unemployment, gangs, suicide, and crimes activities. However, foreign-born workers have a lower probability of joining unions. This is important for policy makers to improve their decisions to help legal immigrants to join unions. The government and unions could also provide more job training to improve immigrants' skills, knowledge, and English proficiency. This might help immigrants to distribute broadly occupations since they are segregated in particular occupation to date.

Table 1.1: Descriptive Statistics for Nonunion and Union Sectors

Explanatory Variable	Nonunion Sector	Union Sector
Hourly Wage	11.407	16.074
	(6.94)	(0.469)
Lwage	2.302	2.668
	(0.498)	(0.469)
Foreign-born workers	0.127	0.105
	(0.333)	(0.307)
Age	38.168	42.486
	(12.233)	(10.674)
Male	0.490	0.564
	(0.500)	(0.496)
White	0.841	0.807
	(0.366)	(0.396)
Married, spouse present	0.562	0.654
	(0.496)	(0.476)
Married, spouse absent	0.011	0.010
	(0.107)	(0.100)
Separated	0.022	0.023
	(0.147)	(0.150)
Divorced	0.106	0.117
	(0.308)	(0.322)
Widowed	0.014	0.015
	(0.119)	(0.120)
Never married/single	0.284	0.181
	(0.451)	(0.385)
Potential experience	18.676	22.620
	(12.187)	(10.846)
No school completed	0.002	0.001
	(0.039)	(0.024)
1st-4th grade	0.004	0.002
	(0.065)	(0.050)
5st-8th grade	0.023	0.014
	(0.151)	(0.119)
9th grade	0.016	0.009
	(0.127)	(0.095)
10th grade	0.029	0.014
	(0.168)	(0.116)
11th grade	0.037	0.018
	(0.188)	(0.135)
12th grade,no diploma	0.012	0.008
	(0.109)	(0.092)
High school grad. or GED	0.304	0.320
	(0.460)	(0.467)
Some college, no degree	0.206	0.194
	(0.404)	(0.395)
Associate degree, occupation program	0.050	0.053
	(0.218)	(0.224)
Associate degree, academic program	0.043	0.045
	(0.204)	(0.206)
Bachelors degree	0.190	0.185
	(0.392)	(0.388)
Masters degree	0.059	0.117

Continued on Next Page...

Table 1.1 – Continued

Explanatory Variable	Nonunion Sector	Union Sector
	(0.236)	(0.321)
Professional degree	0.013	0.009
	(0.112)	(0.094)
Doctorate degree	0.012	0.011
	(0.108)	(0.104)
Northeast	0.205	0.280
	(0.403)	(0.449)
Midwest	0.249	0.280
	(0.432)	(0.449)
South	0.315	0.186
	(0.464)	(0.389)
West	0.232	0.254
	(0.422)	(0.435)
Number of children	0.896	1.036
	(1.137)	(1.151)
Health problem	0.021	0.020
	(0.145)	(0.140)
FirmSize1(under 10)	0.145	0.036
	(0.352)	(0.186)
FirmSize2(10-24)	0.108	0.037
	(0.310)	(0.188)
FirmSize3(25-99)	0.146	0.088
	(0.353)	(0.284)
FirmSize4(100-499)	0.153	0.163
	(0.360)	(0.369)
FirmSize5(500-999)	0.060	0.080
	(0.237)	(0.272)
FirmSize6 (1000+)	0.388	0.596
	(0.487)	(0.491)
Private sector	0.882	0.536
	(0.322)	(0.499)
Full time job	0.827	0.918
	(0.378)	(0.274)
Live in metro	0.779	0.811
	(0.415)	(0.391)
No pension plan at work	0.417	0.156
	(0.493)	(0.363)
Pension plan at work, but not included	0.130	0.071
	(0.336)	(0.256)
Included in pension plan at work	0.453	0.773
	(0.498)	(0.419)
Observations	102,619	21,031

Note: Sample includes all workers between 15 to 65 years of age. Standard errors are in parentheses.

Table 1.2: Descriptive Statistics for Native- and Foreign-Born Workers

Explanatory Variable	Native-Born	Foreign-Born
Hourly Wage	10.923	9.230
	(7.968)	(7.617)
Lwage	2.372	2.296
	(0.516)	(0.474)
Union	0.158	0.131
	(0.365)	(0.338)
Age	38.687	38.231
	(13.553)	(14.404)
Male	0.481	0.491
	(0.500)	(0.500)
White	0.846	-
	(0.361)	(-)
Married, spouse present	0.550	0.596
	(0.498)	(0.491)
Married, spouse absent	0.008	0.035
	(0.089)	(0.183)
Separated	0.021	0.032
	(0.144)	(0.176)
Divorced	0.103	0.063
	(0.303)	(0.243)
Widowed	0.018	0.018
	(0.133)	(0.132)
Never married/single	0.300	0.257
	(0.458)	(0.437)
Potential experience	19.400	20.469
	(13.351)	(13.028)
No school completed	0.001	0.014
	(0.036)	(0.119)
1st-4th grade	0.002	0.041
	(0.040)	(0.197)
5st-8th grade	0.018	0.142
	(0.135)	(0.349)
9th grade	0.027	0.054
	(0.163)	(0.227)
10th grade	0.047	0.039
	(0.211)	(0.193)
11th grade	0.051	0.038
	(0.219)	(0.192)
12th grade,no diploma	0.014	0.027
	(0.116)	(0.161)
High school grad. or GED	0.312	0.247
	(0.463)	(0.431)
Some college, no degree	0.205	0.127
	(0.403)	(0.332)
Associate degree, occupation program	0.045	0.026
	(0.208)	(0.159)
Associate degree, academic program	0.039	0.026
	(0.194)	(0.160)
Bachelors degree	0.164	0.142
	(0.370)	(0.349)
Masters degree	0.054	0.050

Continued on Next Page...

Table 1.2 – Continued

Explanatory Variable	Native-Born	Foreign-Born
	(0.227)	(0.219)
Professional degree	0.012	0.014
	(0.110)	(0.117)
Doctorate degree	0.009	0.014
	(0.094)	(0.118)
Northeast	0.202	0.243
	(0.401)	(0.429)
Midwest	0.251	0.116
	(0.434)	(0.320)
South	0.310	0.264
	(0.463)	(0.441)
West	0.237	0.377
	(0.425)	(0.485)
Number of children	0.877	1.153
	(1.153)	(1.310)
Health problem	0.077	0.050
	(0.267)	(0.218)
FirmSize1(under 10)	0.201	0.229
	(0.400)	(0.420)
FirmSize2(10-24)	0.093	0.116
	(0.291)	(0.321)
FirmSize3(25-99)	0.127	0.148
	(0.333)	(0.355)
FirmSize4(100-499)	0.135	0.137
	(0.342)	(0.344)
FirmSize5(500-999)	0.055	0.051
	(0.228)	(0.220)
FirmSize6 (1000+)	0.389	0.318
	(0.488)	(0.466)
Private sector	0.815	0.904
	(0.388)	(0.295)
Full time job	0.820	0.867
	(0.384)	(0.339)
Live in metro	0.752	0.927
	(0.431)	(0.259)
No pension plan at work	0.431	0.608
	(0.495)	(0.488)
Pension plan at work, but not included	0.118	0.094
	(0.323)	(0.292)
Included in pension plan at work	0.451	0.297
	(0.498)	(0.457)
Observations	107,899	15,751

Note: Sample includes all workers between 15 to 65 years of age. Standard errors are in parentheses.

Table 1.3: The Probability of Joining Unions for Foreign-Born Workers

Variable	Coefficient	Standard Errors	Observations
All Foreign-born Workers	-0.099***	(0.016)	212114
Male Foreign-born Workers	-0.118***	(0.022)	106901
Female Foreign-born Workers	-0.063***	(0.025)	104284
Foreign-born Workers in the Public Sectors	-0.155***	(0.036)	36702
Foreign-born Workers in Private Sectors	-0.055***	(0.019)	175014
Foreign-born Workers in East Region	0.052*	(0.029)	45394
Foreign-born Workers in MidWest Region	-0.210***	(0.042)	53958
Foreign-born Workers in South Region	-0.106***	(0.038)	61386
Foreign-born Workers in West Region	-0.125***	(0.031)	48691

Note: Sample includes all workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. All observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations. Row 1 is the probit regression including all foreign-born workers. Row 2 includes only male foreign-born workers. Row 3 includes only female foreign-born workers. Row 5 is the probit regression including those foreign-born workers in the public sector. Row 6 is the probit regression including those foreign-born workers in the public sector. Row 8 to row 11 are four separate probit regressions for those foreign-born workers in four different regions.

Table 1.4: Estimated Log Wage Equations for Foreign- and Native-Born Workers

Explanatory Variable	Nonunion	Union	Native-Born	Foreign-Born
Im	-0.051***	-0.059***		
	(0.004)	(0.010)		
Union			0.133***	0.113***
			(0.004)	(0.011)
Male	0.101***	0.108***	0.103***	0.091***
	(0.003)	(0.008)	(0.003)	(0.008)
White	0.031***	0.078***	0.053***	-
	(0.004)	(0.008)	(0.004)	(-)
Married, spouse present	0.053***	0.022***	0.048***	0.049***
	(0.004)	(0.009)	(0.004)	(0.010)
Married, spouse absent	0.004	0.048*	0.036***	-0.031*
	(0.012)	(0.028)	(0.015)	(0.018)
Separated	0.014*	0.020	0.006	0.041***
	(0.008)	(0.017)	(0.008)	(0.018)
Divorced	0.032***	0.044***	0.033***	0.044***
	(0.005)	(0.011)	(0.005)	(0.016)
Widowed	0.022***	0.024	0.020***	0.021
	(0.010)	(0.025)	(0.010)	(0.026)
Potential experience	0.015***	0.018***	0.016***	0.011***
	(0.000)	(0.001)	(0.000)	(0.001)
Potential experience squ.	-0.000***	-0.000***	-0.000***	-0.000***
	(9.04e-06)	(0.000)	(9.20e-06)	(0.000)
1st-4th grade	-0.038	0.092	-0.024	-0.003
	(0.032)	(0.111)	(0.121)	(0.026)
5st-8th grade	-0.010	0.063	-0.035	0.012
	(0.029)	(0.102)	(0.109)	(0.024)
9th grade	0.040	0.059	-0.013	0.071***
	(0.030)	(0.104)	(0.109)	(0.026)
10th grade	0.069***	0.074	0.021	0.078***
	(0.030)	(0.103)	(0.109)	(0.027)
11th grade	0.078***	0.085	0.032	0.063***
	(0.030)	(0.102)	(0.109)	(0.028)
12th grade,no diploma	0.073***	0.089	0.028	0.072***
	(0.030)	(0.104)	(0.109)	(0.028)
High school grad. or GED	0.139***	0.185*	0.095	0.140***
	(0.029)	(0.101)	(0.109)	(0.025)
Some college, no degree	0.166***	0.222***	0.126	0.159***
	(0.029)	(0.101)	(0.109)	(0.026)
Associate degree, occupation program	0.207***	0.256***	0.164	0.221***
	(0.030)	(0.101)	(0.109)	(0.032)
Associate degree, academic program	0.207***	0.250***	0.166	0.180***
	(0.030)	(0.102)	(0.109)	(0.032)
Bachelors degree	0.323***	0.340***	0.283***	0.271***
	(0.030)	(0.101)	(0.109)	(0.027)
Masters degree	0.438***	0.472***	0.416***	0.333***
	(0.032)	(0.104)	(0.109)	(0.036)
Professional degree	0.402***	0.318***	0.380***	0.262***
	(0.042)	(0.132)	(0.113)	(0.067)
Doctorate degree	0.474***	0.728***	0.553***	0.182***
	(0.060)	(0.133)	(0.122)	(0.089)

Continued on Next Page...

Table 1.4 – Continued

Explanatory Variable	Nonunion	Union	Native-Born	Foreign-Born
Midwest	-0.077***	-0.028	-0.114***	-0.143***
	(0.012)	(0.037)	(0.012)	(0.061)
South	-0.097***	-0.097***	-0.111***	-0.080
	(0.013)	(0.034)	(0.013)	(0.068)
West	0.172***	0.027	0.172***	-0.210***
	(0.014)	(0.040)	(0.014)	(0.096)
Number of children	0.003*	0.009***	0.006***	-0.004
	(0.001)	(0.003)	(0.001)	(0.003)
Health problem	-0.082***	-0.049***	-0.082***	-0.027
	(0.009)	(0.017)	(0.008)	(0.028)
FirmSize2(10-24)	-0.007	-0.003	-0.005	0.002
	(0.005)	(0.021)	(0.005)	(0.012)
FirmSize3(25-99)	-0.001	0.029	0.002	0.022*
	(0.005)	(0.018)	(0.005)	(0.012)
FirmSize4(100-499)	-0.000	0.016	-0.003	0.028***
	(0.005)	(0.017)	(0.005)	(0.012)
FirmSize5(500-999)	0.011*	0.053***	0.013***	0.035***
	(0.006)	(0.019)	(0.006)	(0.016)
FirmSize6 (1000+)	0.024***	0.085***	0.030***	0.060***
	(0.005)	(0.017)	(0.005)	(0.012)
Private sector	-0.001	0.060***	0.013***	-0.003
	(0.007)	(0.011)	(0.006)	(0.020)
Full time job	0.073***	0.077***	0.072***	0.080***
	(0.004)	(0.011)	(0.004)	(0.010)
Live in metro	0.057***	0.057***	0.059***	0.040***
	(0.003)	(0.007)	(0.003)	(0.013)
Pension plan at work,	0.016***	0.009	0.009***	0.062***
but not included	(0.004)	(0.012)	(0.004)	(0.011)
Included in pension plan	0.117***	0.116***	0.116***	0.128***
at work	(0.003)	(0.008)	(0.003)	(0.009)
State Dummies	YES	YES	YES	YES
Industry Dummies (144)	YES	YES	YES	YES
Occupation Dummies (229)	YES	YES	YES	YES
Year Dummies	YES	YES	YES	YES
Observations	102619	21031	107899	15751
Adj-R2	0.5523	0.4844	0.5762	0.5192

Note: Sample includes all workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. The estimates are from four separate OLS regressions. All observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations.

Table 1.5: Estimated Log Wage Equations for Foreign- and Native-Born Workers by Gender

Explanatory Variable	Male			
	Nonunion	Union	Native-	Foreign-
Im	-0.064***	-0.062***		
	(0.006)	(0.014)		
Union			0.139***	0.112***
			(0.005)	(0.015)
Obs	46231	13277	50867	8641
Adj-R2	0.5349	0.4425	0.5555	0.5215
Chi2(1)	0.03		3.01	
Prob>chi2	0.8708		0.0827	
Chi2(45)	3227.14		3945.99	
Prob>chi2	0.0000		0.0000	
Explanatory Variable	**Female**			
	Nonunion	Union	Native-	Foreign-
Im	-0.033***	-0.056***		
	(0.006)	(0.016)		
Union			0.115***	0.097***
			(0.006)	(0.015)
Obs	56388	7754	57032	7110
Adj-R2	0.5712	0.5554	0.5868	0.5547
Chi2(1)	1.81		1.44	
Prob>chi2	0.1783		0.2307	
Chi2(45)	11448.00		9315.08	
Prob>chi2	0.0000		0.0000	

Note: Sample includes workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. The estimates are from four separate OLS regressions for both male and female workers. All observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations.

Table 1.6: Estimated Log Wage Equations for Foreign-Born Workers from Different Birthplace I

Explanatory Variable	Mexican-Born Workers			
	Nonunion	Union	Native-	Foreign-
Im	-0.119***	-0.059***		
	(0.008)	(0.0235)		
Union			0.133***	0.224***
			(0.004)	(0.023)
Obs	93757	19265	108537	4485
Adj-R2	0.5608	0.4851	0.5757	0.5299
Chi2(1)	6.22		16.22	
Prob>chi2	0.0126		0.0001	
Chi2(120)	5315.69		6848.70	
Prob>chi2	0.0000		0.0000	

Explanatory Variable	Asian-Born Workers			
	Nonunion	Union	Native-	Foreign-
Im	-0.026***	-0.041***		
	(0.009)	(0.021)		
Union			0.133***	0.074***
			(0.004)	(0.024)
Obs	92861	19322	108537	3646
Adj-R2	0.5610	0.4864	0.5757	0.5975
Chi2(1)	0.49		6.73	
Prob>chi2	0.4851		0.0095	
Chi2(45)	4218.00		9319.06	
Prob>chi2	0.0000		0.0000	

Explanatory Variable	Foreign-Born (South America)			
	Nonunion	Union	Native-	Foreign-
ImSa	-0.085***	-0.072***		
	(0.014)	(0.029)		
Union			0.133***	0.135***
			(0.004)	(0.047)
Obs	90596	18945	108537	1004
Adj-R2	0.5626	0.4852	0.5757	0.6401
Chi2(1)	0.15		0.00	
Prob>chi2	0.7001		0.9470	
Chi2(45)	4678.06		12344.89	
Prob>chi2	0.0000		0.0000	

Note: Sample includes workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. The estimates are from four separate OLS regressions for each group of foreign-born workers. All observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations.

Table 1.7: Estimated Log Wage Equations for Foreign-Born Workers from Different Birthplace II

Explanatory Variable	European-Born Workers			
	Nonunion	Union	Native-	Foreign-
Im	-0.013	-0.058***		
	(0.011)	(0.024)		
Union			0.133***	0.063***
			(0.004)	(0.029)
Obs	91411	19160	108537	2034
Adj-R2	0.5620	0.4843	0.5757	0.6026
Chi2(1)	3.05		6.88	
Prob>chi2	0.0808		0.0087	
Chi2(45)	5670.16		8713.24	
Prob>chi2	0.0000		0.0000	

Explanatory Variable	African-Born Workers			
	Nonunion	Union	Native-	Foreign-
Im	-0.040	-0.162***		
	(0.025)	(0.049)		
Union			0.133***	-0.002
			(0.004)	(0.085)
Obs	90075	18820	108537	358
Adj-R2	0.5629	0.4859	0.5757	0.8838
Chi2(1)	4.86		6.49	
Prob>chi2	0.0275		0.0109	
Chi2(45)	965.69		34637.93	
Prob>chi2	0.0000		0.0000	

Note: Sample includes workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. The estimates are from four separate OLS regressions for each group of foreign-born workers. All observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations.

Table 1.8: Estimated Log Wage Equations for Foreign- and Native-Born Workers by Private and Public Sectors

Explanatory Variable	Private Sector			
	Nonunion	Union	Native-	Foreign-
Im	-0.052***	-0.071***		
	(0.005)	(0.012)		
Union			0.142***	0.119***
			(0.004)	(0.011)
Obs	92806	14147	92468	14485
Adj-R2	0.5550	0.5209	0.5822	0.5134
Chi2(1)	2.27		3.55	
Prob>chi2	0.1317		0.0596	
Chi2(45)	4915.37		7830.59	
Prob>chi2	0.0000		0.0000	

Explanatory Variable	Public Sector			
	Nonunion	Union	Native-	Foreign-
Im	-0.038***	-0.023		
	(0.018)	(0.020)		
Union			0.100***	0.078***
			(0.008)	(0.028)
Obs	9813	6884	15431	1266
Adj-R2	0.5516	0.4482	0.5327	0.6435
Chi2(1)	0.29		0.69	
Prob>chi2	0.5896		0.4068	
Chi2(45)	14321.44		15619.88	
Prob>chi2	0.0000		0.0000	

Note: Sample includes workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. The estimates are from four separate OLS regressions for each sector. All observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations.

Table 1.9: Estimated Log Wage Equations for Native- and Foreign-Born Workers by Region

Explanatory	Native-Born Workers			
Variable	East	Midwest	South	West
Union	0.113***	0.132***	0.129***	0.146***
	(0.008)	(0.006)	(0.008)	(0.009)
Obs	21397	31938	30986	23578
Adj-R2	0.5608	0.5959	0.5804	0.5823
Explanatory	**Foreign-Born Workers**			
Variable	East	Midwest	South	West
Union	0.062***	0.137***	0.067***	0.143***
	(0.019)	(0.027)	(0.025)	(0.019)
Obs	3746	2019	3916	6070
Adj-R2	0.4881	0.6254	0.5601	0.5818
Chi2(1)	6.80	0.03	5.92	0.02
Prob>chi2	0.0091	0.8682	0.0150	0.8952
Chi2(77)	10403.65	7794.62	10297.73	5149.55
Prob>chi2	0.0000	0.0000	0.0000	0.0000

Note: Sample includes workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. The estimates are from four separate OLS regressions for both native- and foreign-born workers. All observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations.

Table 1.10: Estimated Log Wage Equations for Native- and Mexican-Born Workers by Region

Explanatory	Native-Born Workers			
Variable	East	Midwest	South	West
Union	0.113***	0.132***	0.129***	0.146***
	(0.008)	(0.006)	(0.008)	(0.009)
Obs	21734	31993	31158	23652
Adj-R2	0.5593	0.5957	0.5802	0.5821
Explanatory	Mexican-Born Worker			
Variable	East	Midwest	South	West
Union	0.330***	0.251***	-0.035	0.243***
	(0.155)	(0.061)	(0.058)	(0.028)
Obs	171	562	1126	2626
Adj-R2	0.8991	0.6714	0.5317	0.6131
Chi2(1)	6.63	5.66	9.95	11.97
Prob>chi2	0.0100	0.0174	0.0016	0.0005
Chi2(77)	27509.76	12259.63	15043.96	5893.35
Prob>chi2	0.0000	0.0000	0.0000	0.0000

Note: Sample includes workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. The estimates are from four separate OLS regressions for both native- and Mexican-born workers. All observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations.

Table 1.11: Estimated Log Wage Equations for Native- and Asian-Born Workers by Region

Explanatory Variable	Native-Born Workers East	Midwest	South	West
Union	0.113***	0.132***	0.129***	0.146***
	(0.008)	(0.006)	(0.008)	(0.009)
Obs	21734	31993	31158	23652
Adj-R2	0.5593	0.5957	0.5802	0.5821
Explanatory Variable	**Asian-Born Workers East**	**Midwest**	**South**	**West**
Union	-0.048	0.046	0.132*	0.075***
	(0.076)	(0.060)	(0.077)	(0.033)
Obs	612	525	597	1912
Adj-R2	0.6624	0.8132	0.8080	0.6121
Chi2(1)	6.81	3.52	0.00	5.05
Prob>chi2	0.0090	0.0607	0.9601	0.0246
Chi2(45)	14362.13	11630.50	14186.33	8374.55
Prob>chi2	0.0000	0.0000	0.0000	0.0000

Note: Sample includes workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. The estimates are from four separate OLS regressions for both native- and Asian-born workers. All observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations.

Chapter 2

The Union Impact on Wages for Foreign-Born Workers: Estimates Using Propensity Score Matching

2.1 Introduction

According to the U.S. Bureau of Labor Statistics, the share of foreign-born workers increased from 14% to 17% between 1994 and 2008. At the same time, the number of foreign-born union members also rose, from 9% to 11%. Since almost one-fifth of US workers are foreign-born, and one-tenth of these foreign-born workers are union members, recently labor researchers are interested in how immigrants affect natives in the labor market and how immigrants are doing in the labor market. This study focuses on how immigrants are affected by unions.

There exists a large literature on the union wage impact across different types of workers. Lewis (1986)'s influential review of the literature on this topic led him to conclude that there is very little difference in the union relative wage effect between female and male workers. Moreover, he argued that it was not clear whether there was a difference in the union relative wage effect between white and nonwhite workers. To date, however, there is no study that has estimated the impact of unions on wages of foreign-born workers versus native-born workers. This paper is the first study on the union wage impact for foreign-born workers and also the first study to compare the union wage premium for foreign-born workers with that for native-born workers. We also explore the union wage premiums for those foreign-born workers who are from different birthplaces and have different skills.

On average, foreign-born workers are less educated, especially Mexican-born workers. Capps et al. (2003) noted that most foreign-born workers, and especially male immigrants, come from Mexico and Asia and work in low-wage jobs. The wages of some of these foreign-born workers are even much lower than the minimum wage. Previous studies suggest that union members and other workers covered by collective agreements receive union wage premiums about 15 percent over nonunion members in the United States. Joining unions could be a good approach for foreign-born workers to receive higher wages. In this paper, our samples include only male foreign-born and male native-born workers to reduce some bias. It would be also interesting to examine the union impact on foreign-born workers who are from Mexico and Asia and who have different skills.

Most studies address the selectivity bias issue and it reduces the reliability of the estimates of the OLS models. However, in addition to bias due to non-random selection, some researchers have argued that if the true OLS model contains higher-order terms or interaction terms between the various variables, there will be additional omitted variable bias in the estimated equation. In addition, various studies have argued that the OLS models estimate the nonoverlapping regions and ignore the common support condition, where there is positive density (overlapping) in the two groups (union and nonunion sectors) being compared. This may lead to additional bias. Therefore, to overcome these shortcomings of the OLS models, previous studies have tried to use propensity score matching (PSM) techniques, which are proposed by Rosenbaum and Rubin (1983). Eren (2007) concludes that "PSM may be considered as robust alternative to OLS for estimation of the average union wage gap". Therefore, in this paper, we estimate union wage premiums using propensity score matching techniques. This paper is the first study on the union wage impact for foreign-born workers using propensity score matching techniques.

In light of the numerous criticisms leveled against estimating the wage differential between union and nonunion workers using the ordinary least squares method, this paper estimates the union impact on the wages of foreign-born and native-born workers using the propensity score matching methodologies and compare the results to those obtained using the OLS approach. From my results, both the propensity score matching and OLS estimates indicate that the union/nonunion wage differentials for male foreign-born workers lie between 12% (OLS) and 27% (PSM). In addition, the difference in terms of the union wage impact between foreign-born and native-born workers is very small (3%). However, foreign-born workers from Mexico have a larger union relative wage effect than those from Asia, and highly skilled Mexican-born workers have the greatest union wage premiums among the three groups that differ in terms of skills.

The remainder of this paper is organized as follows. Section **??** reviews the literature, and Section 3.3 introduces the empirical model. Section 2.4 describes the data, Section 2.5 presents the results, Section 2.6 provides the conclusion, and Section **??** describes the future works.

2.2 Literature Review

The impact of unions on wages has been the focus of a substantial number of papers. However, there has still been controversy over the technique used to examine the union/nonunion wage differentials.

2.2.1 Ordinary Least Squares Regression

There has been substantial disagreement over the estimation of the union wage impact by using ordinary least squares regression (OLS). (Freeman and Medoff, 1984; Lewis, 1986; Aidt and Tzannatos, 2002; Hirsch, 2004). Most studies address the selectivity bias issue and it reduces the reliability of the estimates of the OLS models. If workers randomly selected into the unionized and nonunionized sectors, we would be able to ignore the omitted variables on the right-hand side of the log wage equation. Then, the estimates of the OLS model would be reliable. However, workers do not randomly choose between unionized or nonunionized sectors in reality. Those omitted variables (ability, motivation and production) that we usually cannot obtain from the observational data sets might be correlated with the union variable. Hence, the estimation of the union/nonunion wage differentials using OLS model might be biased. Researchers have argued that this selection process causes a bias to the union wage impact (Card, 1996; Hirsch and Schumacher, 1998; Budd and Na, 2000; Hildreth, 2000). In addition, since union employers have to pay higher wages than nonunion ones, they may choose workers who exhibit higher quality among job applicants. Therefore, this selection process overestimates the union/nonunion wage differentials (Lewis, 1986; Farber, 2001; Bryson, 2002).

Eren (2007) mentioned that most researchers included each possible variable (with the square of experience, square of age) on the right-hand side of the wage equation. If the true OLS model contains higher-order terms or interaction terms between the various variables, there would be omitted variable bias. The second shortcoming of the OLS models that Eren addressed is that these methods extrapolate over nonoverlapping regions, but do not extrapolate only over overlapping regions which exhibit positive density in the two groups compared (union/nonunion). This may lead to another source of bias. Moreover, Farber (2001) observed:"The standard cross-sectional regression technique can be interpreted as the average difference in wages between union and nonunion workers, but it cannot be interpreted as the effect of union membership on the wage of a particular worker."

Because of these controversies, some researchers have tried to use other alternative methodologies to examine the union/nonunion wage differentials. In his survey of the U.S. literature regarding the union relative wage effects, Lewis (1986) referred to two other kinds of methodologies that some researchers had tried to apply to improve the estimates of the OLS models. One of them was the simultaneous equations approach, which uses an equation determining union status (OLS, logit, or probit) and wage equations simultaneously. The second one is the panel data approach. Some studies have also found that problems arise with the case of panel data. For instance, the observations regarding workers who switch from the nonunionized sectors to the unionized sectors are really few. Moreover, there is a strong assumption that the probability of switching between these two sectors is not related to omitted variables for panel data studies (Farber, 2001). Card (1996) addressed another problem by stating that "longitudinal estimators are highly sensitive to measurement error; even a small fraction of misclassified union status changes can lead to significant biases if the true rate of mobility between union and nonunion jobs is low." Therefore, the estimates of panel data studies are usually smaller than those of the OLS models in terms of measurement error bias. After surveying the problems with the above two methodologies and comparing the results with those of the OLS models, Lewis (1986) concluded that the estimates based on the simultaneous equations and panel data approaches were less reliable than those obtained using the OLS method.

2.2.2 The Propensity Score Matching Method

Propensity score matching, proposed by Rosenbaum and Rubin (1983), is a semi-parametric statistical matching approach which tries to randomly match the control units (nonunion workers) as closely as possible with the treatment units (union workers), and then computes the mean wage impact of those matched units given the propensity score. As a result, the main difference between two matched units is only the union variable. Based on this idea, Becker and Ichino (2002) pointed out that propensity score matching is a way of "correcting" the estimation of treatment effects while controlling for the existence of these unobservable variables.

Moreover, propensity score matching methods can impose a common support region by dropping those control observations whose propensity scores are lower than the minimum or those observations whose propensity scores are higher than the maximum. Since these methods can effectively compare treated and control groups over overlapping regions, the problems with the OLS models that estimate over nonoverlapping regions can also be overcome. Furthermore, Bryson et al. (2002) address another advantage of matching in that functional form assumptions are not required for the outcome equation. They argued that the OLS methods impose a functional form between the left-hand-side and right-hand-side variables which may or may not be accurate and which propensity score matching methods avoid. In their paper, which examined the effectiveness of propensity score matching methods, Rosenbaum and Rubin (1983) concluded that propensity score matching can reduce 90% of the bias due to these two specification concerns, and other researchers have also confirmed their results (Foster, 2003; Park and Saccomanno, 2007; Solivas et al., 2007; Chen and Zeiser, 2008). Solivas et al. (2007) conclude that propensity score matching is one way of reducing the selection bias and their paper showed that the propensity score matching was successful in reducing the bias on the covariates.

Because of the above advantages of the propensity score matching methods, such approaches have become increasingly popular in the evaluation of economic policy (Heckman et al., 1997, 1998a; Dehejia and Wahba, 1999, 2002; Bryson, 2002; Bryson et al., 2002; Becker and Ichino, 2002; Eren, 2007). Dehejia and Wahba (1999), using National Supported Work (NSW), Current Population Survey (CPS) and Population Survey of Income Dynamics (PSID) data, found estimates of the treatment impact that are close to the benchmark unbiased estimate obtained from the experiment when using propensity score matching. Their paper is based on Lalonde (1986)'s study on the comparison between experimental and nonexperimental methods used to estimate the causal effect. In addition, Dehejia and Wahba (2002) further contribute to the literature by using different matching methods such as nearest neighbor and radius (caliper) and confirmed that propensity score matching methods are able to yield reasonably accurate estimates of the treatment impact.

So far, there are only two researchers who have used the propensity score matching method to estimate union/nonunion wage differentials. Bryson (2002) used British data from the Workplace Employee Relation Survey of 1998 and found that the raw union/nonunion wage differentials were in the range of 18-25%, depending on the different subgroups. After using propensity score matching, the union/nonunion wage differential falls to 3.5%, and is statistically insignificant. He concluded that "the higher pay of unionized workers is largely accounted for by their better underlying earnings capacity, which is associated with their individual characteristics, the job they do, and the workplaces they find themselves in." However, Bryson (2002) did not compare the estimates of the propensity score matching method with those of the OLS model. Eren (2007) used Panel Study of Income Dynamics data in 1993 to examine the union impact on wages for the private sector in the United States. He employed both the OLS and propensity score matching methods and found that the estimate of the OLS model (24.2%) underestimated the union impact compared with the estimates based on the propensity score matching methods (ranging from 31% to 33%). He also found that omitted variables had a smaller impact on matching results estimates.

2.3 Empirical Model

2.3.1 Estimating the Union/Nonunion Wage Differentials with Propensity Score Matching

Propensity Score

Rosenbaum and Rubin (1983) proposed the propensity score matching methodology in 1983 and defined the propensity score, $p(X)$, as the conditional probability of receiving a treatment given a set of observed covariates.

$$p(X) \equiv Pr\{U = 1|X\} = E\{U|X\} \tag{2.1}$$

where $U = 1, 0$ is an indicator for receiving the treatment (unionized sector) or not receiving the treatment (nonunionized sector). X is a set of observed covariates.

For a given propensity score, we can estimate the average treatment effect on the treated (ATT), which is the union/nonunion wage differential in this paper. ATT is the mean effect of treatment on those who receive treatment compared to those who do not receive treatment given the propensity score,

$$
\begin{aligned}
ATT &\equiv E\{Y_1 - Y_0 | U = 1, X\} \\
&= E\{E\{Y_1 - Y_0 | U = 1, p(X)\}\} \\
&= E\{E\{Y_1 | U = 1, p(X)\} - E\{Y_0 | U = 0, p(X) | U = 1\}
\end{aligned} \tag{2.2}
$$

where Y_1 and Y_0 are log hourly wages (potential outcomes) in the unionized sector (treatment group) and nonunionized sector (control group), respectively.

For the propensity score matching method, there are two fundamental assumptions:

Assumption 1: For a given propensity score ($p(X)$), the set of observed covariates is balanced. In other words, a set of observed covariates is independent of a union variable with the same propensity score.

$$U \quad \perp \quad X \quad | \quad p(X) \tag{2.3}$$

Assumption 2: Unconfoundedness is given the propensity score:

$$Y_1, Y_0 \quad \perp \quad U \quad | \quad X \tag{2.4}$$

$$Y_1, Y_0 \quad \perp \quad U \quad | \quad p(X) \tag{2.5}$$

Rosenbaum and Rubin (1983) pointed out that "if receiving the treatment is random within cells defined by X, it is also random within cells defined by the values of the mono-dimensional variable $p(X)$". Therefore, the potential outcomes are also independent of union variables conditional upon the same propensity score $p(X)$.

In sum, if being in the unionized sector is random, treatment (unionized sector) and control groups (nonunionized sector) should be identically averaged after giving the propensity score (Chen and Zeiser, 2008). Eren (2007) mentioned that matching is a powerful methodology because it can solve the first two bias problems, which Heckman et al. (1998a) addressed in their paper. These are the bias due to a lack of sufficient overlap in the two groups and the bias due to differences in the distributions of the Xs under the common region. Both of these problems are sometimes found to occur in OLS models.

Matching with Propensity Score

The two most common matching methods used to estimate ATT, given the propensity scores, are Nearest Neighbor Matching and Kernel Matching.

In Nearest Neighbor Matching, a treatment unit is matched to a control unit with the nearest propensity score. T and C denote the treatment and control sets. Y_i^T and Y_j^C refer to log hourly wages of the treatment and control units. $C(i)$ denotes the set of control units that are matched to the treatment units given the propensity score ($p(X_i)$),

$$C(i) = \min_j \|p(X_i) - p(X_j)\| \tag{2.6}$$

The average treatment effect on the treated (ATT) is

$$ATT^N = \frac{1}{N^T} \sum_{i \in T} \{Y_i^T - Y_j^C\} \tag{2.7}$$

where N_T is the number of treated units and T denotes all treated observations.

In Kernel Matching, the outcome of a treated unit is matched to a weighted average of the outcomes of all control units.

$$ATT^K = \frac{1}{N^T} \sum_{i \in T} \left[Y_i^T - \sum_{j \in C} g_{ij} Y_j^C \right] \tag{2.8}$$

where g_{ij} is the weight.

According to Becker and Ichino (2002)'s paper, propensity score matching methods only reduce, but do not eliminate, the bias from omitted variables. The bias can only be fully eliminated if joining the unionized sector is truly random among workers who have the same propensity score. They also point out that there is no best propensity score matching method and they also describe some pitfalls for each matching method. For instance, the nearest neighbor matching method tries to match all treated units to control units with the nearest propensity score. Some of these matches might be poor because the nearest control units might have matches of low quality.

2.3.2 Sensitivity Analysis for Average treatment Effects on the Treated

Since propensity score matching has become increasingly popular to evaluate treatment effects, checking the sensitivity of estimated treatment effects on the treated has become an important topic for researchers lately. It is important to know what happens to the estimated results when there are deviations from the underlying identifying conditional independence assumption.

Model

According to Becker and Caliendo (2007), they assume that the participation probability is given by $P_i = P(x_i, u_i) = P(D_i = 1, x_i, u_i) = F(\beta x_i + \gamma u_i)$, where x_i are the observed variables for individual i, u_i is the unobserved variable, and γ is the effect on the participation decision. If there is no unobserved bias, γ will be zero. The probability of receiving treatment will only be determined by x_i. However, if there is unobserved bias, two individuals with the same observed variable x have different probability of receiving treatment. The model assumes that a matched pair of individuals i and j and F is the logistic distribution. The odds that individuals receive treatment are then given by $P_i/1 - P_i$ and $P_j/1 - P_j$, and the odds ratio is given by

$$\frac{\frac{P_i}{1-P_i}}{\frac{P_j}{1-P_j}} = \frac{P_i(1-P_j)}{P_j(1-P_i)} = \frac{exp(\beta x_i + \gamma u_i)}{exp(\beta x_j + \gamma u_j)} \tag{2.9}$$

If both individuals have identical observed variables (x_i), the x vector cancels out, then the odds ratio becomes

$$\frac{exp(\beta x_i + \gamma u_i)}{exp(\beta x_j + \gamma u_j)} = exp\left\{\gamma(u_i - u_j)\right\} \tag{2.10}$$

If there are no differences in unobserved variables ($u_i = u_j$), the odds ratio is one which means there is no unobserved selection bias. Likewise, if unobserved variables have no influence on the probability of receiving treatment ($\gamma = 0$), the odds ratio is also equal to one. Sensitivity analysis now evaluates different γ and $u_i - u_j$ to find out how they alter the estimated treatment effects. Becker and Caliendo (2007) follow Aakvik (2001) and assume that the unobserved covariate is a dummy variable with $u \in \{0, 1\}$. Rosenbaum (2002) shows that (2.9) implies the following bounds on the odds ratio that either of the two matched individuals will receive treatment:

$$\frac{1}{e^{\gamma}} \leq \frac{P_i(1-P_j)}{P_j(1-P_i} \leq e^{\gamma} \tag{2.11}$$

When $e^{\gamma} = 1$, both matched individuals have the same probability of receiving treatment. Otherwise, if for example $e^{\gamma} = 2$, individuals who appear to be similar (in terms of x) could differ in their odds of receiving the treatment by as much as a factor of 2. Rosenbaum (2002) determined that e^{γ} is a measure of the degree of departure from a study that is without unobservable bias .

MH test statistic

For binary outcomes, Aakvik (2001) suggests using the Mantel and Haenszel (1959) test statistic. The MH nonparametric test compares the matched individuals in the treatment group and control group with the same expected number. According to Becker and Caliendo

(2007), researchers must make the individuals in the treatment and control groups as similar as possible, because this test is based on random sampling. Rosenbaum (2002) shows that the test statistic Q_{MH} can be bounded by two known distributions. If $e^{\gamma} = 1$ the bounds are equal to the base scenario of no hidden bias. With increasing e^{γ}, the bounds move apart, reflecting uncertainty about the test statistics in the presence of unobserved selection bias. Let Q^{+}_{MH} be the test statistic, given that we have overestimated the treatment effect, and Q^{-}_{MH}, the case where we have underestimated the treatment effect. The two bounds are then given by

$$Q^{+}_{MH} = \frac{\left| Y_1 - \sum_{s=1}^{S} \widetilde{E}^{+}_{s} \right| - 0.5}{\sqrt{\sum_{s=1}^{S} Var(\widetilde{E}^{+}_{s})}} \tag{2.12}$$

$$Q^{-}_{MH} = \frac{\left| Y_1 - \sum_{s=1}^{S} \widetilde{E}^{-}_{s} \right| - 0.5}{\sqrt{\sum_{s=1}^{S} Var(\widetilde{E}^{-}_{s})}} \tag{2.13}$$

where $\tilde{E}_s$ and $Var(\tilde{E}_s)$ are the large-sample approximations to the expectation and variance of the number of successful participants when u is binary and for given γ. y is the outcome for both treated and control groups and s is stratum.

2.4 Data

The data for this analysis are obtained from the IPUMS-CPS (Current Population Survey), a monthly U.S. household survey conducted jointly by the U.S. Census Bureau and the Bureau of Labor Statistics from 1962 to 2008. It offers a large sample size and complete information on union membership and coverage since 1983, and has also included data identifying foreign-born workers since 1994. Therefore, our data are drawn from 1994 to 2008. The foreign-born population includes immigrants, legal non-immigrants (e.g., refugees and persons on student or work visas) and persons illegally residing in the United States. The term "native-born worker" refers to people residing in the United States who are citizens of the United States in one of three categories: (1) people born in one of the 50 states and the District of Columbia; (2) people born in the United States insular areas, such as Puerto Rico or Guam; or (3) people who were born abroad to at least one parent who was a United States citizen.

All data refer to wage and salary workers who are employed and are 16 years of age and older. We only utilize male workers. The dependent variable is log hourly wages and the variable of interest is union membership. The sample consists of 50,867 native-born workers and 8,641 foreign-born workers. It also consists of 46,231 nonunion workers and 13,277 union workers, as well as 3,116 Mexican-born workers and 1,692 Asian-born workers. Our analysis controls for the following covariates: potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 state effects, 15 year effects, 144 industry effects, and 229 occupation effects.

Mean log hourly wages for nonunion workers amount to 2.378 and for union workers 2.740. Mean log hourly wages amount to 2.475 and 2.343 for native-born and foreign-born workers, respectively. In the nonunion sectors, mean log hourly wages amount to 2.393 and 2.294 for native- and foreign-born workers and 2.757 and 2.600 for native- and foreign-born workers in the union sector.On average, union members have more potential experience ($age - education - 6$), are less educated, have more children and are more likely to be older, male, married, non-white, and native-born workers. Moreover, the private sector is more likely to be unionized. On average, foreign-born workers have more potential experience, are less educated, have more children and are more likely to be younger and married. Furthermore, foreign-born workers are more likely to remain in the private sector, metropolitan areas, and nonunionized sector.

2.5 Results

Table 2.1 presents both the OLS and propensity score matching estimates of union wage premiums for foreign-born and native-born workers. The OLS results show that the union/nonunion wage differential for male foreign-born workers is a positive value of 0.112 (12 percent), whereas for male native-born workers it is 13.9 log points (15 percent). The percentage union/nonunion wage differential is obtained from $(exp(\hat{\delta}) - 1)$. Both values are statistically significant. These estimates are consistent with the conventional wisdom union and nonunion wage gap (Lewis, 1986; Aidt and Tzannatos, 2002). Our results show that the impact of unions on native-born workers' wages is 3-percentage points greater than that on foreign-born workers' wages and is statistically significant. A Chow test rejects the null hypothesis of equality of all the coefficients between foreign-born and native-born workers.

A common support condition is imposed by propensity score matching to improve the quality of the matches. We present results based on nearest-neighbor matching and kernel matching using the Epanechnikov kernel with a bandwidth of 0.63 which are utilized by Eren (2007). Nearest-neighbor matching indicates a positive value of 0.150 (16 percent) union wage premiums for foreign-born workers. Similarly, kernel matching estimate indicates 0.239 (27 percent) union wage premiums for foreign-born workers. For native-born workers, the estimate based on nearest neighbor matching is 0.180 (20 percent) and the result of kernel matching is 0.255 (29 percent). All estimates are statistically significant. Again, both propensity score matching results indicate that the impact of unions on native-born workers' wages is higher than that on foreign-born workers' wages. Moreover, both the nearest neighbor matching and kernel matching results are higher than the OLS results for both native-born and foreign-born workers. These results are similar to Eren (2007) who concludes that the OLS estimates underestimate the union wage impact.

Table 2.2 presents the OLS and propensity score matching estimates of union wage premiums for those foreign-born workers from Mexico and Asia. Our OLS results show that the union/nonunion wage differential for male foreign-born workers who are from Mexican is 23.1 log points (26 percent), whereas for male Asian-born workers it is -0.013 and statistically insignificant. Nearest-neighbor matching indicates a positive value of 0.279 (32 percent) union wage premiums for Mexican-born workers. Similarly, kernel matching estimate indicates 0.349 (42 percent) union wage premiums for Mexican-born workers. Moreover, Asian-born workers give rise to opposite results. The estimate obtained from the OLS is statistically insignificant. However, the results of propensity score matching are statistically significant but smaller than any other groups.

Table 2.3 presents the OLS results of union wage impact for three different skilled groups. Foreign-born workers who are less skilled (with less than a high school degree) have the highest union/nonunion wage differential (17.2 log points; 19 percent). This is consistent with the previous literature in that the union wage impact is larger for less-skilled workers in countries such as Canada, Mexico, the United kingdom, and the United States (Lewis, 1986; Freeman and Medoff, 1984; Simpson, 1985; Fairris, 2003). Native-born workers who have high school degree and higher have larger union/nonunion wage differentials. Our results are consistent with Moll (1993) who finds that unskilled nonwhite workers have the highest union wage premium, but semiskilled and skilled white workers have higher union wage premium. Within this immigrant group, however, the Mexican-born workers tell a different story that the union wage premium is relatively greater (39.1 log points; 48 percent) for the highly skilled workers (college degree or higher). All the estimates based on skill levels for the Asian group are statistically insignificant.

Table 2.4 to table 2.6 present both the OLS and propensity score matching results for three different skilled groups. We only estimate union wage premiums for all native- and all foreign-born workers. When we decompose the data into three different skill groups, we don't have enough matched units for Mexican- and Asian-born workers. For those foreign-born workers who don't have high school degree, they have higher union wage premiums (18.8-25.5 percent) than native except the estimates of the kernel matching. However, those foreign-born workers who have high school degree and higher have lower union wage premiums than native-born workers who have the same education. Again, the results from nearest neighbor and kernel matching are higher than those from OLS for both native-born and foreign-born workers. Our results suggest that the OLS estimates underestimate the union wage impact for both natives and immigrants.

2.6 Conclusions

This paper examines the union wage premiums using Current Population Survey from 1994 to 2008. The methodologies we use in this paper to examine the union/nonunion wage differentials for both foreign-born and native-born workers are the OLS model and propensity score matching method. Our results indicate that the union/nonunion wage differentials for male foreign-born workers are 12% (OLS) and 27% (PSM), depending on the different methods adopted. For male native-born workers, the union/nonunion wage differentials are 15% (OLS) and 29% (PSM). Our evidence suggests that there are few differences in terms of union wage premiums between native-born and foreign-born workers. Also, our results confirm previous literature that the estimate obtained by the OLS model underestimate the union relative wage effect compared with the estimates derived by the propensity score matching methods.

Since immigrant group has been getting larger and larger in the United States, policy makers and researchers have to pay more attention on these documented and undocumented immigrants to know how they are doing in the labor market. Our results show that immigrants do gain much more earnings when they are in the union or being paid by collective agreement. Unions do help immigrants to earn more. A better earning capability of immigrants would likely result in a reduction of social problems such as unemployment, gangs, suicide, and crimes activities from previous studies. How to help and encourage immigrants to join unions to increase their wages in order to decrease social problems will be an important decision for policy makers.

Table 2.1: OLS/Matching Estimates of Union Impact on Wages I

Methodology	Native-Born Workers		
	UNWD	N. Treat.	N. Control
OLS	0.139***		
	(0.005)		
Nearest Neighbor Matching	0.180***	11849	38612
	(0.006)		
Kernel Matching	0.255***	11849	38612
	(0.005)		

Methodology	Foreign-Born Workers		
	UNWD	N. Treat.	N. Control
OLS	0.112***		
	(0.015)		
Nearest Neighbor Matching	0.150***	1416	6686
	(0.017)		
Kernel Matching	0.239***	1416	6686
	(0.014)		

Note: Sample includes male workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. For OLS and PSM, all observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations. The OLS observations for male native-born workers are 50,867 and those for male foreign-born workers are 8,641. UNWD stands for the union/nonunion wage differential.

Table 2.2: OLS/Matching Estimates of Union Impact on Wages II

Methodology	Mexican-Born Workers		
	UNWD	N. Treat.	N. Control
OLS	0.231***		
	(0.031)		
Nearest Neighbor Matching	0.279***	332	1943
	(0.035)		
Kernel Matching	0.349***	332	1943
	(0.029)		

Methodology	Asian-Born Workers		
	UNWD	N. Treat.	N. Control
OLS	-0.013		
	(0.038)		
Nearest Neighbor Matching	0.135***	271	866
	(0.044)		
Kernel Matching	0.131***	271	866
	(0.040)		

Note: Sample includes male workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. For OLS and PSM, all observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations. The OLS observations for male Mexican-born workers are 3,116 and those for male Asian-born workers are 1,692. UNWD stands for union/nonunion wage differential.

Table 2.3: Estimated Log Wage Equations for Male Foreign-Born and Native-Born Workers by Skill

Explanatory Variable	High School Dropout			
	Native-Born	Foreign-Born	Mexican-Born	Asian-Born
UNWD	0.127***	0.172***	0.231***	0.013
	(0.016)	(0.026)	(0.040)	(0.152)
Obs	6976	3399	2044	265
Adj-R2	0.5639	0.4762	0.5179	0.8446

Explanatory Variable	High School Graduate			
	Native-Born	Foreign-Born	Mexican-Born	Asian-Born
UNWD	0.136***	0.096***	0.258***	0.021
	(0.007)	(0.028)	(0.062)	(0.098)
Obs	21489	2528	722	512
Adj-R2	0.5021	0.5019	0.6956	0.6448

Explanatory Variable	College Degree and Higher			
	Native-Born	Foreign-Born	Mexican-Born	Asian-Born
UNWD	0.135***	0.074***	0.391***	-0.014
	(0.008)	(0.027)	(0.096)	(0.052)
Obs	22402	2714	350	915
Adj-R2	0.5515	0.5759	0.8653	0.6947

Note: Sample includes male workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. The estimates are from four separate OLS regressions for each skilled group. All observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations. UNWD stands for the union/nonunion wage differential.

Table 2.4: OLS/Matching Estimates of Union Wage Premiums for High School Dropouts

Methodology	**Native-Born Workers**		
	UNWD	N. Treat.	N. Control
OLS	0.127***		
	(0.016)		
Nearest Neighbor Matching	0.137***	821	5290
	(0.021)		
Kernel Matching	0.252***	821	5290
	(0.017)		
Methodology	**Foreign-Born Workers**		
	UNWD	N. Treat.	N. Control
OLS	0.172***		
	(0.026)		
Nearest Neighbor Matching	0.171***	387	2354
	(0.029)		
Kernel Matching	0.227***	387	2354
	(0.023)		

Note: Sample includes male workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. For OLS and PSM, all observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations. The OLS observations for male native-born workers are 6,976 and those for male foreign-born workers are 3,399. UNWD stands for the union/nonunion wage differential.

Table 2.5: OLS/Matching Estimates of Union Wage Premiums for High School Graduates

Methodology	Native-Born Workers		
	UNWD	N. Treat.	N. Control
OLS	0.136***		
	(0.007)		
Nearest Neighbor Matching	0.179***	5401	15766
	(0.007)		
Kernel Matching	0.240***	5401	15766
	(0.007)		
Methodology	**Foreign-Born Workers**		
	UNWD	N. Treat.	N. Control
OLS	0.096***		
	(0.028)		
Nearest Neighbor Matching	0.154***	454	1614
	(0.030)		
Kernel Matching	0.213***	454	1614
	(0.025)		

Note: Sample includes male workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. For OLS and PSM, all observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations. The OLS observations for male native-born workers are 21,489 and those for male foreign-born workers are 2,528. UNWD stands for the union/nonunion wage differential.

Table 2.6: OLS/Matching Estimates of Union Wage Premiums for Workers with College Degree and Higher

Methodology	**Native-Born Workers**		
	UNWD	N. Treat.	N. Control
OLS	0.135***		
	(0.008)		
Nearest Neighbor Matching	0.175***	5594	16396
	(0.009)		
Kernel Matching	0.226***	5594	16396
	(0.008)		
Methodology	**Foreign-Born Workers**		
	UNWD	N. Treat.	N. Control
OLS	0.074***		
	(0.027)		
Nearest Neighbor Matching	0.133***	509	1616
	(0.032)		
Kernel Matching	0.166***	509	1616
	(0.027)		

Note: Sample includes male workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. For OLS and PSM, all observable covariates includes potential experience, potential experience squared, number of children, as well as dummies for race, marital status, education, region, size of firm, class of worker, pension plan at work, 51 states, 15 years, 144 industries, and 229 occupations. The OLS observations for male native-born workers are 22,402 and those for male foreign-born workers are 2,714. UNWD stands for the union/nonunion wage differential.

Table 2.7: Sensitivity Analysis for Estimates of Nearest Neighbor Matching for Native-Born Workers

Gamma	$Q_{mh}+$	$Q_{mh}-$	$p_{mh}+$	$p_{mh}-$
1	.	.	.	.
1.05	-.006496	-.006496	.502591	.502591
1.1	-.006496	-.006496	.502591	.502591
1.15	-.006496	-.006496	.502591	.502591
1.2	-.006496	.	.502591	.
1.25	-.006496	-.006496	.502591	.502591
1.3	.	.	.	.
1.35	-.006496	-.006496	.502591	.502591
1.4	-.006496	-.006496	.502591	.502591
1.45	-.006496	-.006496	.502591	.502591
1.5	-.006496	-.006496	.502591	.502591

Note: Gamma : odds of differential assignment due to unobserved factors
$Q_{mh}+$: Mantel-Haenszel statistic (assumption: overestimation of treatment effect)
$Q_{mh}-$: Mantel-Haenszel statistic (assumption: underestimation of treatment effect)
$p_{mh}+$: significance level (assumption: overestimation of treatment effect)
$p_{mh}-$: significance level (assumption: underestimation of treatment effect)

Chapter 3

The Determinants and Effects of Job Training on U.S. Immigrant Workers

with Charles Mutsalklisana

3.1 Introduction

In recent years, there have been a growing number of studies on job training and program evaluation in the United States. Despite the large and steadily increasing population of immigrants working in the United States, very few training studies focus on immigrants. "Immigrants make up one in nine U.S. residents, one in seven U.S. workers, and one in five low-wage workers" (Capps et al., 2003). Since the year 2000, legal immigrants that came to the U.S. have numbered approximately one million per year (Borjas, 2008). Surprisingly, only modest knowledge of the effect of training programs on minorities in the labor market is known, and the evidence of these programs on immigrants as a whole is even more limited (Ashenfelter, 1983; Capps et al., 2003; Flores-Lagunes et al., 2007). This paper aims to improve our understanding of the effect of job training on immigrant workers.

The main incentives of job training provided by employers is not only to increase worker's productivity, but also to minimize the adoption time when new technology arrives. These improvements are in line with business goals, including implementation in order to reduce costs of job turnover. It is widely known in the labor literature that firms which provide job training have more productive workers and less expenses associated with turnover (Frazis et al., 1998). These firms can earn more profit from the increase of workers' productivity. Thus, it is likely that companies would want to invest in their employees, particularly immigrants who tend to possess fewer skills, limited English proficiency and less formal education; hence, they are opportunities for productivity gains. Immigrants also need training to help them adapt to changing work environments. However, our results suggest that immigrant workers receive considerably less training than native workers–21.6 percent and 39.1 percent respectively.

Job training is an important instrument that can be used to assist both native and immigrant workers improve their welfare. Most immigrants, like other low-skill workers, lack opportunities to learn new skills and to benefit from employer-provided training programs (Ahlstrand et al., 2001). In the private sector, it is often the case that not all employees receive equal training opportunities. More specifically, employers prefer to provide job training to workers who are more likely to stay with the firms for a relatively long period of time. With evidences from the US and Canada showing that workers are not likely to pay for their own training (Parent, 1999), immigrants with greatest need of general training are often unable or unwilling to invest in training themselves. Hence, these immigrants, who are low-skilled and are perceived to have short-term tenure, are in need of general training, and yet they end up losing out on opportunities to be trained.

In the case of government sponsored training, these programs often have stringent requirements that reduce immigrants' chances of attaining job training. Some studies have found that low-wage immigrants are often under-served due to the lack of proficiency in English (Tumlin and Zimmermann, 2003). The main purpose of government training programs is to prepare economically disadvantaged individuals, such as welfare program recipients for the labor market, yet it is common to find certain minimum requirements that prevent low-skilled workers from obtaining these training opportunities (Capps et al., 2003). For instance, it is frequently observed that programs require participants to have at least ninth-grade level literacy, and they should have numeracy ability and basic English skills.

According to the March 2002 Current Population Survey (CPS), 18 percent of all immigrant workers and 28 percent of all low-wage immigrant workers have educational attainment less than ninth grade, while 1 percent of all native workers and 2 percent of all low-wage native workers have educational attainment less than ninth grade (Capps et al., 2003). Thus, by not meeting the minimum requirements, many immigrants lose out on the opportunity to be trained by the programs that are sponsored by government.

Immigrant workers are important assets in the U.S. labor market, representing 14 percent of the total U.S. labor force and 20 percent of low-wage earning workers (Capps et al., 2003). Immigrant workers represent a large share of low-skilled workers, who engage in low paying jobs that are necessary for our economy. Another important fact is that nearly half of immigrants earn less than two times the minimum wage, compared to less than one-third for native workers (Capps et al., 2003).

Policy makers also might be interested in the labor market outcomes of immigrants besides earnings. Immigrants face not only low wage, but also problems of low education, limited English proficiency and lack of formal training. At the same time, the immigrant workforce is also confronted by basic problems such as inadequate healthcare, transportation and childcare (Edid, 2007). A better earning capability of immigrants would likely result in a reduction of social problems such as unemployment, gangs and crimes activities. Non-experimental data have shown that increased earnings reduce criminal activity (Lalonde, 1986) and lower the number of crimes such as murders (Donohue and Siegelman, 1998).

To summarize, we know that training is important to immigrants. Yet, immigrants, as a minority group who is in need of training, are receiving much lesser training than natives. Hence, to help policy makers improve their decisions regarding training provided for immigrant workers, we need to gain better understanding of the effect of training on immigrants.

The contribution of this paper is that it is one of the few papers that inquires into the effect of training on immigrants, and it is the first paper that proposes an economic model that analyzes the effect of job training on immigrant workers in the United States. This is the first paper that looks at both the mean and distributional effect of job training on immigrant workers in the United States. It is the first paper that studies the effect of training on immigrant workers using the 2004, 2001 and 1996 Survey of Income and Program Participation (SIPP) data, and we do this using a Quantile Regression (QREQ) model, a semi-parametric reweighting DiNardo, Fortin and Lemieux (DFL) method, and propensity score matching method. Another contribution is that this is one of the few studies to explore the effect of training on native born workers excluding foreign born workers.

This paper: (1) compares the average impact of job training on earnings of native and immigrant workers; (2) explores the distributional impacts of job training on native and immigrant workers' earnings and (3) examines the counterfactual distributional impact of job training if trained and untrained workers have similar observable characteristics.

3.1.1 Issues on Training

In general, there are two main problems when measuring the impact of training: unobserved heterogeneity and selection. First, an unobservable heterogeneity problem occurs when the model suffers from omitted variables bias. In our study, the omitted variables come from the difficulty of measuring characteristics such as intelligence, motivation and obtaining the opportunity cost of participating in training.[1]

Since the focus of our paper is to study the comparative effect of training on natives and immigrants, obtaining the exact estimate is not our biggest problem. There might exist upward bias in the estimated training coefficients, yet it will not change our result. Since, we are looking at the result of the comparison which is the differences in return to training between natives and immigrants.

Since this paper is the first paper on the topic, we believe it is more important to layout the fundamental results rather than getting into a sophisticated econometric model.

[1] Like other studies on training, we acknowledge that we will not be able to resolve all unobserved heterogeneity problems.

Therefore, the other issue of training evaluation such as selection problem is beyond the scope of this paper. Yet, we will discuss how to overcome some of the selection problems. Also, we plan to tackle the selection issue in our future work.

Lastly, there is an issue of the importance of the distributional effects of training. It is known that the effects of training often vary across the wage distribution that are not captured by the mean; examples of these studies include Lalonde (1986) or Abadie et al. (2002). The distributional outcomes beyond simple averages are of fundamental interest in the policy analysis of welfare implications such as transfer, education and training programs (Abadie et al., 2002). The results from the distributional analysis highlight the important differences for low-wage and low-skill workers.

The paper is structured as follows. Section 3.2 reviews the literature on job training. Section 3.4 describes the data used in our analysis. Section 3.3 explores the methods that will be applied. Section 3.5 discusses our analysis and findings. Section 3.6 summarizes our findings and comments on policy implications.

3.2 Literature Review

Despite the fact that there are many studies exploring the different aspects of immigrant workers in the United States and numerous studies measuring the effect of training individually, we know of only a few studies that inquire into the impact of training on immigrants. We suspect that the reasons for the small number of studies could be twofold: because of poor data sources and because there are very few training programs that target immigrants specifically. Few existing studies that inquire into the impact of training on immigrants present only basic results such as composition tables and graphs. For instance, Capps et al. (2003) compare the profile of low-wage immigrant workers and native workers using composition tables and graphs on CPS data. They argue in favor of revamping training requirements to increase access to training for immigrant workers. Another source such as

Edid (2007) evaluates the training needs of immigrant workers in Syracuse by using interviews and composition graphs on Census data. She supports the improvement of English proficiency among among other skills needed for immigrant workers.

Fortinand and Parent (2005) explore the incidence of training in US and Canada. Although the paper does not focus on immigrants, the authors note that immigrant status is significant in the determinants of training. They also found that immigrants in the U.S. are more likely to have shorter training than those in Canada. Though these papers did not evaluate the effect of training on immigrants, they provide the cornerstone question for us to explore further. We will improve upon their studies by setting up our training evaluation model, comparing the differences in the returns to training for immigrant and native workers, doing so using econometric tools. In the rest of our literature review, we will discuss the general training aspects[2] and the relevant econometric evaluation methods.

Most studies on the effect of job training are from the United States and Europe. The majority of the studies are empirical works. They focus on questions such as (1) do participants benefit from job training, (2) is there social merit in job training or (3) what are the determinants of job training? The general consensus is that government sponsored job training programs are ineffective, resulting in a small positive or even negative net benefits, yet with great heterogeneity (Heckman et al., 1999). Furthermore, it appears that there is no consensus on the model to measure the effect of job training program. Due to the inconsistencies of measurement findings, there is an increase in econometric methodological studies in recent years. Generally, the results from experimental studies yield impact of job training programs on earnings that range from minus 15 to plus 70 percent (Heckman et al., 1999) [3].

The focus of our paper is on non-experimental data, focusing on the impact of training of disadvantage workers in the United States[4]. Notable studies include economically

[2]This is not specifically for immigrants, since this is one of the very first papers on this topic

[3]Experimental studies are training programs that design to have randomization of program participants largely composing of government sponsored programs.

[4]Majority of the non-experiment job training studies used data from Comprehensive Employment and Training Act, 1982 (CETA), Manpower Development and Training Act, 1962 (MDTA), Trade Adjustment Assistance Program (TAA) and Job Training Partnership Act, 1982 (JPTA).

disadvantaged adult participants (Ashenfelter and Card, 1985; Bassi, 1983; Dickinson et al., 1986), displaced workers (Bloom, 1990; Decker and Corson, 1995) and economically disadvantaged youth (Dickinson et al., 1986; Bryant and Rupp, 1987; Bassi et al., 1984).

Besides the studies that focus on displaced workers and youths, there are many papers that explore the effect of training on specific minority groups. For example, there are papers that focus on the effect of training on blacks (Butler and Heckman, 1977; Smith and Welch, 1986; Kane, 1994; Flores-Lagunes et al., 2007), and the effect of training on Hispanic workers (Schochet et al., 2001; Flores-Lagunes et al., 2006, 2007). These literatures lead us to believe that there are heterogeneous treatment effects in these special groups, since women, disabled, youth and other minority groups have different treatment effects. Although these papers above discuss the effect of training on different races, the studies on immigrants, which is a diverse minority group that has different characteristics than natives, have not yet been fully investigated as we mentioned before.Below, we discuss the literature of the most relevant approaches to this paper, studying the average effect and the distributional effect.

3.2.1 Study of Average Treatment Effect on Treated

For the average effect of training on workers in the US, the conventional methods use RE or FE models on panel data. Panel data can be used to control for unobserved omitted variables. Furthermore, recent studies estimate the effect of job training using a propensity score matching (PSM) method, first proposed by Rosenbaum and Rubin (1983). It assumes that the distributional outcome of the treated is not statistically different from the distributional outcome of the untreated. PSM resolves selection on combinations and interactions of observable characteristics, but does little to address bias due to unobserved heterogeneity.

There have been many studies using PSM, including the main contribution to the training literature by Dehejia and Wahba (1999, 2002), Heckman et al. (1998b), and Heckman et al. (1997, 1998a). Dehejia and Wahba (1999), using Lalonde (1986) study of National Supported Work (NSW), CPS and PSID data, found a reduction in the treated coefficient estimates of experimental data when using PSM. In addition, Dehejia and Wahba (2002) further contribute to the literature by using several matching methods such as nearest neighbor and radius (caliper). Other main extensions of original PSM include the kernel and stratification matching methods. When choosing among the existing matching methods, the general consensus is that there is no preferred matching method (Becker and Ichino, 2002). Dehejia and Wahba (1999, 2002) concluded that "propensity score matching methods provide a natural weighting scheme that yields unbiased estimates of the treatment impact for nonexperimental approaches."

3.2.2 Study of the Distribution Effect

The effect of job training on the income distribution has become of great interest in recent years. This is due to the importance of the distributional effect of training, which is not captured by the mean impact. The most predominant method is applying the basic framework of quantile regression that was developed by Koenker and Gilbert Bassett (1978). QREQ assumes conditional treatment and no selection problem. The latest distribution studies focus on resolving the treatment (selection problems) on both conditional and unconditional effects, where exogenous treatment choices assume selection on observable and endogenous treatment choices assume selection on unobservable (Frolich and Melly, 2008).

Some recent distributional studies focus on conditional treatment with selection on observables only. Quantile treatment effect (QTE), proposed by Abadie et al. (2002), who study the effect of the JTPA training program using IV estimator on conditional quantiles in order to deal with bias due to unobservable. Using "indicators of the randomized offered training as binary instrument variable" in QTE, they found the largest impact of job training at low quantiles for women and the only positive impact in the upper half distribution of men with JTPA data. As a benchmark, treatment 2SLS estimates a 15 percent increase in earning for women and a 9 percent increase in earning for men (Orr et al., 1996). Unfortunately, our non-experimental data does not possess the IV that was used in Abadie et al. (2002)[5]. Our paper extends the current literature by analyzing the effect of training on immigrants in the United States using quantile regression, reweighting methodology, and propensity score matching method.

3.3 Empirical Model

3.3.1 A Model of Quantile Regression

In this section, we examine the estimator of the Koenker and Gilbert Bassett (1978) quantile regression model. The quantile regression model has outcome variable, Y, binary treatment indicator, D, and a vector of covariates, X. In our empirical study of job training, Y is workers' earnings and D is an indicator of exposure to job training. X indicates the observable characteristics of the workers (occupation sorting, demographic differences and human capital differences). For n observations, individual workers' outcomes can be expressed as follows:

$$Y_i \equiv Y_i^1 D_i + Y_i^0 (1 - D_i) \tag{3.1}$$

where Y_i^1 is the indicator of potential outcome if workers received treatment (potential earning if workers received training) and Y_i^0 is indicator of potential outcome if workers did not receive treatment (potential earning if workers did not receive training) for the entire wage distribution function. Quantile regression model has the following basic assumptions Frolich and Melly (2008):

Assumption (1): Suppose the outcome is a linear function of X and D, the outcome can be expressed as follows:

[5] Similar to our average effect study, we acknowledge that we will not be able to resolve all unobserved heterogeneity problems in our distributional effect study

$$Y_i^d = X_i\beta^t + D\delta^t + \varepsilon_i \tag{3.2}$$

$$Q_\varepsilon^t = 0 \tag{3.3}$$

where Q_ε^t is the t^th quantile. We assume D is uncorrelated with error term ε

Assumption (2): Independence:

(Y^0, Y^1) is jointly independent of $D|X$

Independence assumption indicates that the potential outcomes are not affected by treatment on unobservable.

In QREQ model, we assume selection is exogenous on observable characteristics, $(Y^0, Y^1)|X$. Hence, we assume that workers' earnings for both the treated and the non-treated group are not affected by exposure or self selection to job training conditional on observable characteristics. The classical quantile regression can then be computed with the following formula:

$$(\beta^t, \delta^t) = arg_(\beta, \delta) min \sum \rho_t (Y_i - X_i\beta - D_I\delta) \tag{3.4}$$

3.3.2 Counterfactual Study

In this section, we review the reweighting technique of the DFL model. Using the DFL reweighting technique, we simulate the counterfactual earnings of native (immigrant) workers along the entire distribution, if these workers, who did not receive job training (D=0), have similar observable characteristics as workers that received training (D=1). Hence, we are comparing the earnings of two groups of workers that possess similar observable characteristics, except that one group has training.

Suppose (W, Z, D) is a vector representing each worker, where W indicates earnings of the workers, Z is a vector of worker observable attributes (e.g. occupation, firm estimated size, education, age and metropolitan area) and D is the training indicator (D=1 for received training or D=0 for did not receive training). The probability of workers not received training conditioned on workers' observable characteristics can be estimated using logit or probit model as $f(D=0|Z)$. The "reweighting function," $\Psi(Z)$, is the counterfactual weight of untrained workers that would have prevailed if untrained workers possessed observable characteristics of trained workers:

$$\Psi(Z) = \frac{dF(Z|D=1)}{dF(Z|D=0)} = \frac{f(Z|D=1)}{f(Z|D=0)} = \frac{f(D=1|Z)/f(D=1)}{f(D=0|Z)/f(D=0)} \tag{3.5}$$

The reweighting function can simply be calculated by the product of the sample weight and [p/(1-p)], where p is the predicted probability of being untrained workers conditioning on their observable attributes. The intuition here is that we are making a better comparison group of untrained workers that look more similar to trained workers by using the reweighting function that allocates additional weight to the observations that belong to the minority categories. For example, since immigrants with no training have much lower education than immigrants with training, more weight is allocated to the higher education untrained immigrant workers. Finally, the hypothetical quantile training premium is simply the difference between actual earnings of workers with training and counterfactual earnings of workers without training[6].

3.3.3 Estimating with Propensity Score Matching

Propensity Score

Rosenbaum and Rubin (1983) proposed the propensity score matching methodology in 1983 and defined the propensity score, $p(X)$, as the conditional probability of receiving a treatment given a set of observed covariates.

[6]For the wage gap density methodology see Antecol and Steinberger (2009). For the detail estimate regression and their assumptions see Pagan and Ullah (1999)

$$p(X) \equiv Pr\{D = 1|X\} = E\{D|X\} \tag{3.6}$$

where $D = 1, 0$ is an indicator for receiving the treatment (job training) or not receiving the treatment (no job training). X is a set of observed covariates.

For a given propensity score, we can estimate the average treatment effect on the treated (ATT). ATT is the mean effect of treatment on those who receive treatment compared to those who do not receive treatment given the propensity score,

$$\begin{aligned} ATT &\equiv E\{Y_1 - Y_0|D = 1, X\} \\ &= E\{E\{Y_1 - Y_0|D = 1, p(X)\}\} \\ &= E\{E\{Y_1|D = 1, p(X)\} - E\{Y_0|D = 0, p(X)|D = 1\} \end{aligned} \tag{3.7}$$

where Y_1 and Y_0 are log hourly wages (potential outcomes) in the treatment group and control group, respectively.

For the propensity score matching method, there are two fundamental assumptions:

Assumption 1: For a given propensity score ($p(X)$), the set of observed covariates is balanced. In other words, a set of observed covariates is independent of a training variable with the same propensity score.

$$D \quad \perp \quad X \quad | \quad p(X) \tag{3.8}$$

Assumption 2: Unconfoundedness is given the propensity score:

$$Y_1, Y_0 \quad \perp \quad D \quad | \quad X \tag{3.9}$$

$$Y_1, Y_0 \quad \perp \quad D \quad | \quad p(X) \tag{3.10}$$

Rosenbaum and Rubin (1983) pointed out that "if receiving the treatment is random within cells defined by X, it is also random within cells defined by the values of the mono-dimensional variable $p(X)$". Therefore, the potential outcomes are also independent of training variables conditional upon the same propensity score $p(X)$.

In sum, if receiving the training is random, treatment and control groups should be identically averaged after giving the propensity score (Chen and Zeiser, 2008). Eren (2007) mentioned that matching is a powerful methodology because it can solve the first two bias problems which are the bias due to a lack of sufficient overlap in the two groups and the bias due to differences in the distributions of the Xs under the common region (Heckman et al., 1998a). Both problems are sometimes found to occur in the OLS models.

Matching with Propensity Score

The two most common matching methods used to estimate ATT, given the propensity scores, are Nearest Neighbor Matching and Kernel Matching.

In Nearest Neighbor Matching, a treatment unit is matched to a control unit with the nearest propensity score. T and C denote the treatment and control sets. Y_i^T and Y_j^C refer to log hourly wages of the treatment and control units. $C(i)$ denotes the set of control units that are matched to the treatment units given the propensity score $(p(X_i))$,

$$C(i) = \min_j \|p(X_i) - p(X_j)\| \tag{3.11}$$

The average treatment effect on the treated (ATT) is

$$ATT^N = \frac{1}{N^T} \sum_{i \in T} \{Y_i^T - Y_j^C\} \tag{3.12}$$

where N_T is the number of treated units and T denotes all treated observations.

In Kernel Matching, the outcome of a treated unit is matched to a weighted average of the outcomes of all control units.

$$ATT^K = \frac{1}{N^T} \sum_{i \in T} \Big[Y_i^T - \sum_{j \in C} g_{ij} Y_j^C \Big] \tag{3.13}$$

where g_{ij} is the weight.

According to Becker and Ichino (2002)'s paper, propensity score matching methods only reduce, but do not eliminate, the bias from omitted variables. The bias can only be fully eliminated if receiving the job training is truly random among workers who have the same propensity score.They also point out that there is no best propensity score matching method and they also describe some pitfalls for each matching method. For instance, the nearest neighbor matching method tries to match all treated units to control units with the nearest propensity score. Some of these matches might be poor because the nearest control units might have matches of low quality.

3.3.4 Sensitivity Analysis for Average treatment Effects on the Treated

Since propensity score matching has become increasingly popular to evaluate treatment effects, checking the sensitivity of estimated treatment effects on the treated has become an important topic lately. Researchers are interested in what happens to the estimated results when there are deviations from the underlying identifying conditional independence assumption.

Model

According to Becker and Caliendo (2007), they assume that the participation probability is given by $P_i = P(x_i, u_i) = P(D_i = 1, x_i, u_i) = F(\beta x_i + \gamma u_i)$, where x_i are the observed variables for individual i, u_i is the unobserved variable, and γ is the effect on the participation decision. If there is no unobserved bias, γ will be zero. The probability of receiving treatment will only be determined by x_i. However, if there is unobserved bias, two individuals with the same observed variable x have different probability of receiving treatment. They assume that a matched pair of individuals i and j and F is the logistic distribution. The odds that individuals receive treatment are then given by $P_i/1-P_i$ and $P_j/1-P_j$, and the odds ratio is given by

$$\frac{\frac{P_i}{1-P_i}}{\frac{P_j}{1-P_j}} = \frac{P_i(1-P_j)}{P_j(1-P_i)} = \frac{exp(\beta x_i + \gamma u_i)}{exp(\beta x_j + \gamma u_j)} \tag{3.14}$$

If both individuals have identical observed variables (x_i), the x vector cancels out, then the odds ratio becomes

$$\frac{exp(\beta x_i + \gamma u_i)}{exp(\beta x_j + \gamma u_j)} = exp\left\{\gamma(u_i - u_j)\right\} \tag{3.15}$$

If there are no differences in unobserved variables ($u_i = u_j$), the odds ratio is one which means there is no unobserved selection bias. Likewise, if unobserved variables have no influence on the probability of receiving treatment ($\gamma = 0$), the odds ratio is also equal to one. Sensitivity analysis now evaluates different γ and $u_i - u_j$ to find out how they alter the estimated treatment effects. Becker and Caliendo (2007) follow Aakvik (2001)' paper and assume that the unobserved covariate is a dummy variable with $u \in \{0, 1\}$. Rosenbaum (2002) shows that (3.14) implies the following bounds on the odds ratio that either of the two matched individuals will receive treatment:

$$\frac{1}{e^{\gamma}} \leq \frac{P_i(1-P_j)}{P_j(1-P_i} \leq e^{\gamma} \tag{3.16}$$

When $e^{\gamma} = 1$, both matched individuals have the same probability of receiving treatment. Otherwise, if for example $e^{\gamma} = 2$, individuals who appear to be similar (in terms of x) could differ in their odds of receiving the treatment by as much as a factor of 2. Thus, Rosenbaum (2002) determined that e^{γ} is a measure of the degree of departure from a study that is without unobservable bias .

MH test statistic

For binary outcomes, Aakvik (2001) suggests using the Mantel and Haenszel (1959) test statistic. The MH nonparametric test compares the matched individuals in the treatment group and control group with the same expected number. According to Becker and Caliendo (2007)'s paper, researchers must make the individuals in the treatment and control groups as similar as possible because this test is based on random sampling. Rosenbaum (2002) shows that the test statistic Q_{MH} can be bounded by two known distributions. If $e^{\gamma} = 1$ the bounds are equal to the base scenario of no hidden bias. With increasing e^{γ}, the bounds move apart, reflecting uncertainty about the test statistics in the presence of unobserved selection bias. Let Q^{+}_{MH} be the test statistic, given that we have overestimated the treatment effect, and Q^{-}_{MH}, the case where we have underestimated the treatment effect. The two bounds are then given by

$$Q^{+}_{MH} = \frac{\left| Y_1 - \sum_{s=1}^{S} \widetilde{E}^{+}_{s} \right| - 0.5}{\sqrt{\sum_{s=1}^{S} Var(\widetilde{E}^{+}_{s})}} \tag{3.17}$$

$$Q^{+}_{MH} = \frac{\left| Y_1 - \sum_{s=1}^{S} \widetilde{E}^{-}_{s} \right| - 0.5}{\sqrt{\sum_{s=1}^{S} Var(\widetilde{E}^{-}_{s})}} \tag{3.18}$$

where $\tilde{E}_s$ and $Var(\tilde{E}_s)$ are the large-sample approximations to the expectation and variance of the number of successful participants when u is binary and for given γ. y is the outcome for both treated and control groups and s is stratum.

3.4 Data Source

In our empirical study, we utilize data from the Survey of Income and Program Participation (SIPP) in the years 1996, 2001 and 2004. SIPP, funded by US Census Bureau, collects a variety of information, such as income, labor force participation, types of jobs, program participation and demographic data. A main objective of SIPP is to forecast the cost and evaluate the impact of government and other social programs in the United States. Yet, numerous studies on private sector also use SIPP data. The dataset contains unique information on job training and immigrants, as well as income, other human capital and occupational information. SIPP has abundant observations, unlike many other datasets that have small sample size issues.

The SIPP surveys, which are conducted by personal visits and by telephone interviews, were first administered in 1984. The surveys interview approximately 14,000 to 36,700 households of individuals 15 years of age and older, civilian non-institutionalized, conducting monthly questionnaires. The SIPP, currently containing 12 waves for each year surveyed, is a panel dataset, collecting data once every few years. Each wave includes the "Core" (mainly containing wave standard variables that evaluate economic situation in the US) and the "Topical Modules" (containing different variables depending on the wave). With 12 individual waves for each year that it conducts the survey, SIPP data has more information than other resources, since the survey can ask different sets of questions to interviewees of different waves. Yet, a drawback is that we do not always observe the same person for each set of questions.

The diverse variable availability in SIPP data and its being "rich enough to determine program eligibility" (Heckman et al., 1999) benefit our analysis because only a few databases contain enough information on both job training and immigrant status. Yet, another problem of using SIPP is that variables are sometime removed or transferred to different waves of questionnaires, changing from one year to another. Particularly, only the 2004 data have the English ability variable and job training variable in the same wave.

From SIPP data, we use the combination of first (Core) wave and second (Module Two) wave. Since it is well documented that there are sizeable differences between the impact of job training on men versus women and adults versus youths (Heckman et al., 1999), we use a sample that include all the adult males between 22 to 65 years of age[7]. In addition, we eliminate observations that have hourly earnings of zero dollars or less. The variables that we are most concerned with are wage and job training. The dependent variable wage is the log of workers' earnings per hour, while the job training variable indicates whether the workers had any job training in the last 10 years[8].

The observable covariates used in our wage equation includes year dummies, race (White, Black and Hispanic) dummies, seniority dummies, seniority squared divide by 100, years of education categorical variables, marital status dummies, dummies variable indicating whether the individual lives in the female headed household, having children younger than 18 living in the family dummies, metropolitan, private firm, firm size dummies, dummy variable denoting possession of health insurance, union dummies, state dummies, industry dummies and occupation dummies[9]. To resolve the concerns relating to changes in some of the variables' meanings and their categories, we re-categorize 1996 and 2001 data to match the 2004 definition.

3.4.1 Statistics Summary

Table 3.1 and Table 3.2 provide the summary statistics of the variables used in our analysis. Specifically, Table 3.1 displays the mean values of total native sample (column 3), native sample which have training (column 6) and native sample which did not have training (column 9). Table 3.2 displays similar statistical summaries for immigrants. The summary tables show that natives and immigrants possess different characteristics. Also, the tables show that workers who have training possess different characteristics than workers who do not have training.

[7] Even though we realize that lower incidences of training may occur at later age, older individuals are included in the sample to capture possible retraining of recent technical advancement, such as computer skills.

[8] Unfortunately, we are not able to differentiate between the effect of job training accumulated much earlier or recently and location of training.

[9] We use age as proxy for seniority.

One notable difference is that native workers (at 39.1 percent) received more training in the past ten years than the immigrant workers (at 21.6 percent) on average. Overall, natives earn higher wages than immigrants, while on average, trained immigrants earn marginally higher wages than untrained natives. Thirdly, the average age of natives is almost two years older than immigrants. Since seniority increases the chance of obtaining training, age differences may lead to a small upward bias in the training estimate for natives. Hence, we control for seniority with age variable.

Next, we observe that immigrant workers have accumulated less education than native born workers (26 percent of immigrants have nine years of education or less compared to only 2 percent of natives). Yet, at the higher education level, immigrants and natives have similar education attainment. The education differences in the immigrants may contribute to the heterogeneous effects in the earnings that we will explore in our distributional study. In addition, we observe that trained workers have accumulated more education than not trained workers for both natives and immigrants.

We observe that more natives have female heads of households and more natives have health insurance than their immigrant counterparts. Yet, more immigrants are likely to be married, are likely to have more children and are likely to live in metropolitan cities. Generally, being married, having fewer children and living in metropolitan cities increase the probability of obtaining training. Last, more immigrants work in private sector and smaller size firms than natives. Since private sector and small size firms tend to provide less job training, it is possible that job selection of immigrants contribute to immigrants receiving less job training. Consequently, we note that the above covariates are important and need to be controlled for in our wage equation.

3.5 Analysis

This section explores the effect of job training, comparing native and immigrant workers' wage premium on training. We divide this section into four main subsections: unconditional effect (mean and quantiles), conditional effect (quantile regression), the DFL reweighting counterfactual method and propensity score matching method.

3.5.1 Unconditional Effect

Table 3.3 exhibits results of the unconditional effect of job training on earnings. It shows that though, on average, immigrants earn less than natives, the unconditional training wage premium for immigrants is relatively larger than natives. In terms of monthly earnings, on average, immigrant workers who received training earned 1480 dollars more than immigrant workers who did not receive training, while trained native workers earned 848 dollars more than not trained native workers. The unconditional training wage premium for natives is 17.7 percent, while the unconditional training wage premium for immigrants is 19.7 percent (Table 3.4). Hence, on average, the unconditional wage premium on training for immigrants is 2 percent higher than that for natives.

Figure 3.1 presents the distributional effect. The patterns displayed from the effect of training on the income distribution for immigrant and native workers are quite different from the effect of training at the mean and each other. For the unconditional training premium, the results show that training has largest proportional impact at the upper quantiles for immigrants and largest proportional impact at the middle quantiles for natives. For natives, it shows the differences in results across quantiles where the unconditional wage premium on training increase from lower to median quantiles and slowly decrease from median to upper quantiles. We observe the highest wage gain from training for middle-income workers and the lowest wage gain from training for low-income and upper-income workers. The .50 quantile natives experience earnings increase of around 21.5 percent, while the .05 and .95 quantile natives experience earnings increase of only 12.5 percent.

Unlike natives, the unconditional wage premium on training for immigrants monotonically increases from lower to upper quantiles. Immigrants enjoy highest unconditional wage premium on training at the upper part of the income distribution and receive lowest wage premium at the lower part of the income distribution. The .10 quantile immigrants experience earnings increase of around 8 percent, while the .95 quantile immigrants experience a substantial increase in earnings of 28 percent.

When comparing the training premiums of natives and immigrants across quantiles, it is notable that the unconditional training premium for immigrants is relatively larger than that of natives at the upper part of the income distribution and similar to natives at the lower and middle quantiles (Figure 3.2). While the unconditional training wage premium for natives is more or less constant across the quantiles, the effect of training on immigrants is more remarkable. A possible reason that there are differences in training premium between natives and immigrants at the higher quantiles is that high-skilled immigrants benefit more from training since they started with lower human capital, in terms of language literacy. For instance, an immigrant doctor will obtain a sizeable income raise once receiving training and English ability is good enough to deal with patients, while an immigrant farmer will earn similar wage increase as a native farmer even after receiving training and English skill is improved.

3.5.2 Conditional Effect

OLS Model

Table 3.5 show the results from pooled OLS controlling for race, age, and education. The estimated training coefficients are positive for the pooled OLS model. The conditional training wage premium for natives is 13.9 percent, while the conditional training wage premium for immigrants is 11 percent. Hence, on average, the conditional wage premium on training for immigrants is 3 percent less than that for natives.

Table 3.6 displays the estimated training coefficients using the pooled OLS estimates of the conditional effect of job training on wages when adding more variables on the right hand side. The results from the OLS model suggests that when we include more observable characteristics to wage equation, the effect of training on earnings reduces considerably. Yet, the estimated training coefficients remain positive and significant. Rows 2 to 10 show the training premium conditioning on different covariates (human capital, demographic and occupation sorting) that influence earnings. Row 2 adds the yearly dummies. Row 3 includes racial dummies. It reports that the effect of training drops to 16.7 percent for natives and to 15.4 percent for immigrants when racial dummies are included. The racial dummies have larger influence on estimated training coefficients for immigrants than natives; four percent and one percent reduction respectively.

Row 4 controls for age and education. The magnitudes of estimated coefficients reduce to 11 percent for immigrants and to 13.9 percent for natives. It is noteworthy that almost half of the training effect for immigrants and one-eighth of the training effect for natives are due to the effects of schooling and work experience. Row 5 shows that when controlling for marital status, female head of household and having children under age of 18, we only find marginal changes in the estimated training coefficients. As expected, while being married and having children have positive impact on earnings, living in the female head of household house has negative impact on earnings. Row 6 adds private firm, metropolitan and firm size dummy variables to the wage equation. Surprisingly, we also found only minor impacts from metropolitan and firm size despite both variables being significant and positively correlated with earnings.

Row 7 includes a health insurance variable. Controlling for health insurance, we observe 3 percent reduction in estimated training coefficient for immigrants. Row 8 includes the union dummy variable. It shows that while being in the union increases wage earning around 14 percent, controlling for union status reduces the effects of training by 0.1 percent for natives and 0.3 percent for immigrants. Rows 9 and 10 show that estimated training coefficients stay relatively the same when state-level and industrial dummy variables are included. Row 11 indicates that adding an occupation covariate reduces the effect of training by 1.3 percent for natives and 1.9 percent for immigrants.

Next, we include interaction terms to the wage equation. When we include industry interaction dummy variables to the wage equation, we found a small reduction in the effect of training for natives. For immigrants, we also found a reduction in the effect of training, yet the training variable become statistically insignificant[10]. When we include occupation interaction dummy variables to the wage equation, we found a small increase in the effect of training for natives and again statistically insignificant coefficient for immigrants. In general, adding interaction covariate terms to the wage equation adds very little change to the training coefficients.

Consequently, we conclude that the differences in racial, schooling, experiences and occupation are the most relevant observable covariates that account for the majority of effect of training. After conditioning on all observable covariates, the OLS model indicates that the training premium for immigrants is lesser than natives.

Sensitivity analysis

Education

This section discusses the result from the sensitivity analysis. It is possible that the training variable is correlated with the covariates, causing spurious training coefficients. Due to the small sample size of immigrants, we perform the sensitivity analysis using the OLS model. In general, we perform the robustness check by modifying our original model, testing our model for high school graduate, college graduate, married and non-married groups. We know from the summary table that native college graduates received more training than high school graduates (nearly 10 percent more), and native high school graduates account for the most untrained native workers (36.9 percent) (Table 3.1). Similarly, we find that immigrant college graduates received more training than high school graduates (nearly 6 percent more), and immigrant high school graduates also account for the most untrained immigrant workers (nearly 25.4 percent) (Table 3.2).

[10]Interaction result table is available upon request.

Table 3.7 displays the estimated training coefficients for workers who graduated from high school versus workers who graduated from college, and married versus non-married workers. Row 1 and 2 show that training premium is relatively higher for natives who have bachelor degrees compared to those native who have high school diplomas. Surprisingly, we find the opposite effect for immigrants, particularly negative effect for immigrants with bachelor degrees. However, we note that the estimated training coefficient for immigrants with bachelor degrees is not statistically significant, and the sample size for immigrants in this group is very small.

Comparing natives and immigrants, we still find that natives enjoy higher training premium than immigrants. Training increases natives' earnings around 5.6 percent versus immigrants' earnings around 5.3 percent for high school graduates. For college graduates, training increases natives' earnings around 6.4 percent, while training reduces earnings for immigrants with college degree around 13.4 percent. The negative return on training for these immigrants is likely due to the small sample size of immigrants with college degree. We summarize that for this sensitive analysis, the effect of training for natives is strong, while the effect of training for immigrants is somewhat weak and ambiguous. Furthermore, it is notable that in general, college graduates have more training than high school graduates for both natives and immigrants, and the effect of training for both secondary school and college graduates is robust with natives enjoying higher training premium.

Marital Status

From the summary Table 3.1, we observe that the native married workers received more training than native non-married workers (nearly 9 percent more). For immigrants, non-married workers received slightly more training than married workers (Table 3.2). Similar to education groupings, we study the effect of training for married and non-married using separate wage equations. Comparing married and non-married, row 3 and 4 show that the effect of training on wage is relatively larger for non-married native workers, while the effect of training on wage is much smaller for non-married immigrants (Table 3.7). Comparing natives and immigrants, the training premium enjoyed by married workers is relatively the same for both natives and immigrants, while the training premium enjoyed by non-married workers is relatively the larger for the natives. However, the estimated training coefficient for immigrants is negative and not statistically significant.

For row 5 and 6, we remove health insurance, union, state, industry and occupation dummies from the wage equation for married and non-married workers (Table 3.7). We obtain statistically significant estimated training coefficient for immigrants, yet the main results do not change. Similar to the result from education, the effect of training for natives is strong, while the effect of training for immigrants is somewhat weak and ambiguous. From the result of the sensitivity analysis, we conclude that our original model is relatively robust where natives enjoy higher training premium than immigrants for the majority of the cases.

Looking at the sensitivity test, there is reason to believe that we will have a heterogeneous outcome. For example, within the high school graduate, both married and older cohort, the differences in return to training are negligible, while for the differences within college graduate, which are not married and younger, they are relatively large. In addition, when we study the average effect using econometric tools such as OLS, it is required that the wage density conform to the normal distribution condition. Yet, this may not always hold. Hence, it is important to explore the distributional effects.

Quantile Regression (QREQ)

Since one of our main concerns is the welfare of different income groups, especially, low-wage workers, the study of the distributional effect of training is particularly important. First, we explore the distributional impact of training on native workers using quantile regression. Figure 3.3 presents (solid line) conditional training wage premium of native workers across the quantiles distribution (the top figure), conditioning on race, age and education covariates, (dashed line) unconditional training wage premium of native workers across quantiles (trained workers' mean wage minus untrained workers' mean wage) and (light straight line) estimated OLS coefficient. It shows that the magnitude of estimated training coefficients are considerably lowered across the distribution after we include yearly, race, age and education dummies to the wage equation. The reduction of training premium amplifies at the upper half of the wage distribution, becoming more uniformly distributed with a small dip at the highest quantiles.

Figure 3.4 displays conditional training premium of natives when we include all observable characteristics to the wage equation (the top figure). The training effect on natives reduces more drastically when we include all observable covariates to the wage equation, becoming even more uniformly distributed, almost identical to the OLS line. Nevertheless, it is noticeable that despite the reduction in wage premium, we still observe that training raises wages throughout the quantiles distribution, at an increase of nearly 8 percent. Hence, according to Quantile regression, we find that training raises wage premium for all native workers in the distribution in a similar fashion.

Next, we study the impact of training on immigrant workers across quantiles. Figure 3.3 displays (solid line) training wage premium of immigrant workers across the quantiles distribution conditioning on race, age and education covariates, (dashed line) unconditional training wage premium of immigrants and (light straight line) estimated OLS coefficient (the bottom figure). Similar to native workers, we observe that the conditioned distributional training premium of immigrant workers is reduced considerably. Yet, the reduction is much more apparent than that of natives, especially at the highest quantiles.

When we include all observable characteristics to the wage equation, we observe that the quantiles wage premium for immigrants drop further (the bottom figure). It changes from rapid monotonically increase of unconditional training premium across the quantiles distribution to slow monotonically decrease of conditional training premium (Figure 3.4). Despite the reduction in training premium, we still observe that training increases earnings at the lower and middle quantiles, an increase of nearly 4 percent.

Consequently, according to the QREQ model, low and mid-income immigrants still have lower training premium than natives, while high-income immigrants, who had higher unconditional training premium, now have much lower conditional training premium than natives. It is noteworthy that similar to OLS estimates, the Quantile regression may suffer from upward bias. As mentioned in the methodology section, we concede that we will not be able to resolve unobservable selection problems. Yet, we will explore the counterfactual study alternative. It can be argued that some differences in wage premium between natives and immigrants are due to their observable characteristics. In the next section, we will further explore the training effect on the distribution using the DFL weighting technique.

3.5.3 Counterfactual Study

In the similar spirit of the DFL, this section presents a counterfactual study of job training, simulating the quantile distribution of training premium, supposing both trained and untrained workers have similar observable characteristics. First, using a counterfactual study, we explore the impact of training on native workers. The top of Figure 3.5 and 3.6 present (dashed line) unconditional job training premium of native workers and (solid line) counterfactual training premium of native workers (trained workers' wage minus untrained workers' wage supposes these untrained native workers have similar observable characteristics as trained workers) (the top figure). When we corrected for observable characteristics differences between trained and untrained native workers, we found ambiguous results.

When we remove race, age and education differences between trained and untrained native workers, we observe that the counterfactual premium becomes slightly more uniform. We find that training premium drops marginally at the upper half and increases negligibly at the lower half of the income distribution, reducing training premium at upper quantiles around 2 to 3 percent and increasing training premium at lowest quantiles around 2 percent (the top of Figure 3.5). The results show that suppose high income untrained natives have similar education as trained natives, they would receive lower wage.

When we removed all observable characteristic differences between trained and untrained natives, we observed that the counterfactual training premium changed from being relatively uniform to a slight monotonically increase from lower to upper quantiles (the top of Figure 3.6). We find that counterfactual training premium generally remain unchanged at the lower half of the income distribution. Yet, at the upper half of the income distribution, the counterfactual training premium is surprisingly greater than the unconditional wage premium. Suppose high income untrained natives have similar observable characteristics as trained natives, this result indicates that untrained natives will actually receive even lower wage.

Next, continuing to apply the counterfactual framework, we explore the impact of training on immigrant workers. The bottom of Figure 3.5 displays (dashed line) unconditional training premium of immigrant workers and (solid line) counterfactual training premium of immigrant workers (trained workers' wage minus untrained workers' wage supposes these untrained immigrant workers have similar observable characteristics as trained workers). Unlike the income premium distribution of native workers, we observe that training premium of immigrant workers reduces considerably after corrected for observable characteristics differences between trained and untrained workers.

When we corrected for race, age and education differences between trained and untrained immigrant workers, we found that the counterfactual training premium was reduced considerably, particularly, at the upper half of the income distribution. Although the effect is minimal at lower quantiles, the reduction of immigrants' wage premium is much greater than the natives, dropping the wage premium at .50 Quantile around 8 percent and at .95 Quantile around 13 percent (the bottom of Figure 3.5). Similar to natives, the results show that if high income untrained immigrants have similar education as trained immigrants, they would receive a lower wage, yet in much larger scale.

The bottom of Figure 3.6 displays counterfactual training premium when we removed all observable characteristics differences between trained and untrained immigrants. We find that the counterfactual training premium is still smaller than unconditional training premium, but the effect from correcting for all observable characteristics differences is not as large as for correcting for only race, age and education differences. Yet, we still find a reduction of wage premium, particularly, at the upper half of the income distribution, dropping the wage premium at upper half of the distribution around 5 percent (the bottom of Figure 3.6).

Using the DFL weighting method, we found that after removing all observable characteristics differences between trained and untrained workers, training still increases wage premium for both natives and immigrants throughout the income distribution. Similar to Abadie et al. (2002) that found the impact of training only at the upper half of the income distribution, we observe largest impact of training at upper half of the income distribution for both natives and immigrants. Hence, high income workers still benefit most from training. These results suggest that training premium for highly skilled workers is higher than lower skilled workers. Also, training premium is lower for immigrants than natives for the majority of the income distribution, and training has the smallest effect for very low-skilled and low-wage immigrant workers. Nevertheless, we still find that training increases wage premium of low and middle income workers, including immigrants.

3.5.4 Propensity Score Matching

Table 3.8 presents both the OLS and propensity score matching results of training premiums. The OLS results show that the job training premium for foreign-born workers is a positive value of 0.039, whereas for native-born workers it is 0.076. Our results show that there is 4-percent difference in the job training premium between native and immigrant. A common support condition is imposed by propensity score matching to improve the quality of the matches. We present results based on nearest-neighbor matching and kernel matching using the Epanechnikov kernel with a bandwidth of 0.63 which are utilized by Eren (2007). Nearest-neighbor matching indicates a positive value of 0.063 (training premium) for foreign-born workers. Similarly, kernel matching estimate indicates 0.184 (training premium). For native-born workers, the estimate based on nearest neighbor matching is 0.108 and the result of kernel matching is 0.229. All estimates are statistically significant. Those matching results are higher than the OLS results for both native-born and foreign-born workers, especially the results of kernel matching. Our results suggest that OLS estimates underestimate the training premium.

3.6 Conclusion

In this paper, we study the effect of job training on the US immigrant workers, using the 1996, 2001 and 2004, Survey of Income and Program Participation (SIPP) data. Job training is the essential key for immigrant workers, who often face immense difficulty in the labor market that tends to favor native workers, to improve their standard of living. Training increases life time earning capability of immigrants, which is rewarded in the labor market and helps reduce poverty driven social problems. Since immigrants are important and necessary part of the US labor market and represent a large fraction of the workers, it is important to address and understand the true effect of training on immigrants in the US.

Earlier studies on training rarely look at immigrants, and few studies that look at the effect of training on immigrants utilize economic models, using instead only descriptive and mean table as analytical tools. Hence, this allows us to study different aspects of training and immigrants that have not been explored. As a result, we improve upon prior studies by setting up our training evaluation model, studying the impact of training on both the average and the distributional earning of workers and comparing the differences in the return to training for immigrant and native workers by applying the Quantile regression (QREQ) model, the DiNardo, Fortin and Lemieux (DFL) reweighting methods, and propensity score matching method.

From our mean analysis, we find that training has a positive and significant effect on wages of the average immigrant worker. Looking at the unconditional training premium, our analysis suggests that though natives earn more than immigrants, the training premium for immigrants is relatively larger than natives. In other words, immigrants, who received job training, earn higher wage premium than natives who have received job training.

Our sensitivity analysis results show that our original model is relatively robust with high school graduates, college graduates, married and non-married workers where natives enjoy higher training premium than immigrants for the majority of cases.

From our distribution study, we find that training has a positive effect on wages of immigrant workers for most parts of income distribution. The results suggest that the effect of training across workers income quantiles is relatively different compared to the effect of training at the mean for both immigrants and natives. The differences in the effect of training appear large, interesting and important for welfare consideration when we look at the effect of training across the different quantiles.

Looking at the unconditional training premium, the results show that training increases earnings throughout the quantiles for both immigrants and natives. We observe that immigrants enjoy largest unconditional training premium at the upper part of workers' earning quantiles, and they enjoy lowest unconditional training premium at the lower part of workers' earning quantiles. We find evidence that natives enjoy largest unconditional training premium at the middle of workers' earning quantiles. Comparing natives and immigrants, it is notable that the unconditional training wage premium for immigrants is considerably larger than natives at upper quantiles and similar to natives at the lower and middle quantiles. As a result, we observe a more remarkable unconditional gain from training for wealthy immigrants and less gain for poorer immigrants.

Examining counterfactual study, the DFL reweighting technique shows that after removing all observable characteristics differences between trained and untrained workers, training still increases wage premium for both natives and immigrants throughout the income distribution. Counterfactual training simulates the quantile distribution of training premium, if untrained workers have similar observable characteristics as trained workers. After controlling for all observable characteristics, we observe a sizeable reduction in training premium for immigrants, yet we note a small increase in training premium for natives.

Similar to Abadie et al. (2002) that found an impact of training only at the upper half of the income distribution, we observe the largest proportional impact of training at upper half of the income distribution for both natives and immigrants. Our analysis provides strong evidence for the hypothesis that after corrected for observable characteristics differences between trained and untrained workers, the effect of training is relatively larger for rich natives, much larger for middle income natives and similar for the poor natives and immigrants. Nevertheless, we still find that training increases wage premium of low and middle income workers, including immigrants.

3.6.1 Policy Implication

There are several proposed initiatives that policy makers can take away from this study. The practical lesson is that job training is beneficial and important to the improvement of immigrants' well-being, yet many immigrants are still deprived of these much needed training. Although we did not find the largest impact of training for low-skilled and low-wage immigrant workers, we did find a strong and positive training impact for this group. Hence, these low-skilled and low-wage immigrants should be one of the main target groups of training provision, since they need the most assistance in obtaining training and would greatly benefit from the result of training.

Policy makers can restructure the existing programs to allow easier excess for immigrants such as revamping the Workforce Investment Act by changing the English prerequisite. Also, it is important to concentrate on outreach programs that increase awareness to Limited English Proficiency (LEP) workers regarding the availability of job training. Furthermore, realizing that these poor workers earn their living day by day, tangible assistance such as providing of transportation and childcare arrangements during training may be necessary. In addition, policy makers may consider offering English as a Second Language (ESL) classes and training programs simultaneously to immigrants, focusing on providing English literacy for agricultural workers and providing of English and skills training for manufacturing and service workers.

For immigrants that are unable to participate in the training program immediately, policy makers can allocate funds for job fairs that target immigrants, providing help with filling out applications and language assisted interviews. For private sector, government agencies can redirect some resources to give companies incentive to provide training for immigrants. Tax cuts and funding can be used as incentive tools to encourage firms to grant training to immigrants and managers to promote workforce diversity.

3.6.2 Future Work

There remain many facets of the effect of training on immigrants that have not been explored. Our framework can be extended to study other minority groups within immigrants, particularly concentrating on women, youth and other racial ethnic immigrants such as Black, Hispanic and Asian. It is important to pay attention to these subgroups, especially youth, since they are the future workforce and would provide life time return on social investment.

From the distribution study, our application of the DFL reweighting technique can further be used to identify the observable characteristics differences between trained and untrained workers that are most influential to training premium at different earning quantiles. Hence, this application is useful in assisting policy makers to pinpoint existing problems. Next, since our DFL reweighting analysis relies on the assumptions that treatment selection is based on observable characteristics, it is possible that selection problem may bias our distribution estimate. An instrument variable to use in quantiles treatment effects (QTE), Abadie et al. (2002) can be another possible research avenue.

In the future, we will try to correct the problem and check whether our treatment effect is significant or not. Furthermore, to further resolve unobserved heterogeneity problem, creating a better panel dataset with the focus on immigrants and training would be very beneficial.

In future work, we plan to explore the effect of training on immigrant work hours and employment. In addition, due to the shortcoming of our SIPP data, we cannot identify time since exposure to training, amount of training received and length of training exposure. With other data set, future research should investigate the effect of length of training exposure similar to Flores-Lagunes et al. (2007).

Table 3.1: Mean Values of Natives

Summary	Obs	Native (All)	Standard Errors	Obs	Training	Standard Errors	Obs	No Training	Standard Errors
Log Hourly Wage	24401	2.499	0.456	9519	2.595	0.448	14882	2.437	0.451
Hourly Wage	24401	13.450	6.011	9519	14.731	6.262	14882	12.628	5.695
Monthly Income	24401	2482	2134	9519	2824	2156	14882	2262	2091
Training last 10 year	24401	0.391	0.488	9519	1.000	0.000	14882	0.000	0.000
Training last 1 year	24401	0.199	0.399	9519	0.508	0.500	14882	0.000	0.000
1 Day to 1 Week	934	0.361	0.480	934	0.361	0.480	0		
More than 1 Week	934	0.333	0.471	934	0.333	0.471	0		
Currently in Training	934	0.109	0.312	934	0.109	0.312	0		
White	23896	0.855	0.352	9300	0.881	0.324	14596	0.839	0.368
Black	23896	0.134	0.340	9300	0.107	0.309	14596	0.151	0.358
Hispanic	24401	0.088	0.283	9519	0.067	0.249	14882	0.101	0.302
Asian	23896	0.011	0.104	9300	0.013	0.111	14596	0.010	0.100
Age	24401	38.985	11.343	9519	38.957	10.826	14882	39.003	11.663
Age square	24401	16.485	9.324	9519	16.348	8.882	14882	16.572	9.597
Highest grade < 9	23735	0.024	0.152	9141	0.010	0.097	14594	0.032	0.177
Highest grade < 12	23735	0.093	0.290	9141	0.055	0.228	14594	0.117	0.321
High school diploma	23735	0.416	0.493	9141	0.347	0.476	14594	0.459	0.498
Some college	23735	0.355	0.479	9141	0.438	0.496	14594	0.303	0.459
Bachelor diploma	23735	0.092	0.289	9141	0.121	0.326	14594	0.073	0.261
Master or higher	23735	0.020	0.139	9141	0.029	0.167	14594	0.014	0.119
Married	24401	0.557	0.497	9519	0.599	0.490	14882	0.530	0.499
Female head	24401	0.053	0.223	9519	0.035	0.184	14882	0.064	0.245
Kids 18 years or less	24401	0.389	0.488	9519	0.406	0.491	14882	0.378	0.485
Metropolitan area	24075	0.741	0.438	9388	0.740	0.439	14687	0.741	0.438
25 to 99 employees	24112	0.241	0.428	9409	0.236	0.424	14703	0.244	0.429
100+ employees	24112	0.423	0.494	9409	0.479	0.500	14703	0.387	0.487
Private sector	24134	0.872	0.334	9419	0.826	0.379	14715	0.901	0.298
Public sector	24134	0.128	0.334	9419	0.174	0.379	14715	0.099	0.298
Health Insurance	24401	0.786	0.410	9519	0.857	0.350	14882	0.740	0.439
Union	18542	0.016	0.125	6815	0.024	0.152	11727	0.011	0.105
Employed	24401	0.970	0.171	9519	0.976	0.153	14882	0.966	0.181
Low English	8922	0.013	0.112	3052	0.003	0.058	5870	0.018	0.132

Source: SIPP 1996, 2001 and 2004. Notes: Sample includes all adults male between 22 to 65 years of age. The omitted category for length of training is less than 8 hours of training, age is 22 to 29 years, education is less than first grade and firm size is under 25 employees.

Table 3.2: Mean Values of Immigrants

Summary	Obs	Immigrant (All)	Standard Errors	Obs	Training	Standard Errors	Obs	No Training	Standard Errors
Log Hourly Wage	4486	2.319	0.435	991	2.457	0.456	3495	2.282	0.422
Hourly Wage	4486	11.217	5.323	991	12.939	6.096	3495	10.743	4.988
Monthly Income	4486	1985	1579	991	2439	2118	3495	1860	1369
Training last 10 year	4486	0.216	0.411	991	1.000	0.000	3495	0.000	0.000
Training last 1 year	4486	0.097	0.296	991	0.449	0.498	3495	0.000	0.000
1 Day to 1 Week	117	0.322	0.469	117	0.322	0.469	0		
More than 1 Week	117	0.345	0.477	117	0.345	0.477	0		
Currently in Training	117	0.112	0.316	117	0.112	0.316	0		
White	3972	0.834	0.372	829	0.779	0.415	3143	0.848	0.359
Black	3972	0.109	0.312	829	0.144	0.351	3143	0.100	0.301
Hispanic	4486	0.600	0.490	991	0.417	0.493	3495	0.650	0.477
Asian	3972	0.057	0.232	829	0.077	0.266	3143	0.052	0.222
Age	4486	37.277	10.764	991	38.216	10.492	3495	37.019	10.824
Age square	4486	15.054	8.744	991	15.705	8.556	3495	14.876	8.788
Highest grade < 9	4429	0.260	0.439	972	0.096	0.294	3457	0.305	0.461
Highest grade < 12	4429	0.132	0.338	972	0.097	0.296	3457	0.141	0.348
High school diploma	4429	0.271	0.444	972	0.259	0.438	3457	0.274	0.446
Some college	4429	0.201	0.401	972	0.336	0.473	3457	0.164	0.371
Bachelor diploma	4429	0.090	0.286	972	0.156	0.363	3457	0.072	0.258
Master or higher	4429	0.028	0.164	972	0.049	0.216	3457	0.022	0.147
Married	4486	0.633	0.482	991	0.626	0.484	3495	0.635	0.481
Female head	4486	0.041	0.198	991	0.040	0.196	3495	0.041	0.199
Kids 18 years or less	4486	0.549	0.498	991	0.530	0.499	3495	0.555	0.497
Metropolitan area	4393	0.903	0.296	967	0.906	0.291	3426	0.902	0.298
25 to 99 employees	4438	0.261	0.439	980	0.270	0.444	3458	0.259	0.438
100+ employees	4438	0.350	0.477	980	0.470	0.499	3458	0.317	0.466
Private sector	4441	0.946	0.227	980	0.891	0.312	3461	0.961	0.194
Public sector	4441	0.054	0.227	980	0.109	0.312	3461	0.039	0.194
Health Insurance	4486	0.537	0.499	991	0.727	0.446	3495	0.485	0.500
Union	3784	0.011	0.102	749	0.026	0.160	3035	0.007	0.082
Employed	4486	0.973	0.162	991	0.967	0.179	3495	0.975	0.158
Low English	1694	0.422	0.494	330	0.218	0.413	1364	0.467	0.499

Source: SIPP 1996, 2001 and 2004. Notes: Sample includes all adults male between 22 to 65 years of age. The omitted category for length of training is less than 8 hours of training, age is 22 to 29 years, education is less than first grade and firm size is under 25 employees.

Table 3.3: Ordinary Least Squares Regression Results.

	Native	Immigrant
Training last 10 year	848.357	1480.882
	(40.459)	(101.594)
Constant	3357.008	2671.511
	(27.458)	(43.054)
N Observations	49642	8256

Source: SIPP 1996, 2001 and 2004. Notes: Sample includes all adults male between 22 to 65 years of age. Standard errors are in parentheses.

Table 3.4: OLS Model. Estimate Effect of Training on Earnings. Dependent Variable: Log of Hourly Earnings.

	Native	Immigrant
Log Hourly Wage	Pooled OLS	Pooled OLS
Training last 10 year	0.177	0.197
	(0.006)	(0.017)
Year2001	0.192	0.219
	(0.007)	(0.017)
Year 2004	0.238	0.249
	(0.007)	(0.017)
Constant	2.288	2.102
	(0.006)	(0.012)
N Observations	24401	4486

Source: Data are from SIPP, and include individuals from year 1996, 2001 and 2004. Notes: Sample includes all adults male between 22 to 65 years of age. The omitted category for age is 22 to 29 years and education is less than first grade. Standard errors are in parentheses.

Table 3.5: OLS Model. Estimate Effect of Training on Earnings. Dependent Variable: Log of Hourly Earnings.

	Native	Immigrant
Log Hourly Wage	Pooled OLS	Pooled OLS
Training last 10 year	0.139	0.110
	(0.006)	(0.018)
Year 2001	0.184	0.211
	(0.007)	(0.018)
Year 2004	0.231	0.229
	(0.007)	(0.017)
White	0.058	0.035
	(0.030)	(0.031)
Black	-0.092	-0.046
	(0.031)	(0.036)
Hispanic	-0.059	-0.117
	(0.012)	(0.018)
Age: 30-39	0.221	0.148
	(0.010)	(0.023)
Age: 40-49	0.301	0.172
	(0.017)	(0.040)
Age: 50-65	0.249	0.124
	(0.028)	(0.073)
Age square	0.002	0.005
	(0.001)	(0.003)
Highest grade $<$ 9	0.243	0.194
	(0.125)	(0.039)
Highest grade $<$ 12	0.334	0.243
	(0.124)	(0.041)
High school diploma	0.434	0.293
	(0.124)	(0.040)
Some college	0.476	0.347
	(0.124)	(0.042)
Bachelor diploma	0.507	0.387
	(0.124)	(0.049)
Master or higher	0.625	0.431
	(0.126)	(0.064)
Constant	1.604	1.719
	(0.127)	(0.056)
N Observations	23896	3972

Source: Data are from SIPP, and include individuals from year 1996, 2001 and 2004. Notes: Sample includes all adults male between 22 to 65 years of age. The omitted category for age is 22 to 29 years and education is less than first grade. Standard errors are in parentheses.

Table 3.6: OLS Models. Estimate Effect of Training on Earnings. Dependent Variable: Log of Hourly Earnings.

Training	Native	Im.	Diff.	Num. (Natives)	Num (Im.)	Firm Size	Health Ins.	Union	State	Ind.	Occ.
1	0.158	0.175	0.017	24401	4486	No	No	No	No	No	No
	(0.006)	(0.017)									
2	0.177	0.197	0.020	24401	4486	No	No	No	No	No	No
	(0.006)	(0.017)									
3	0.167	0.154	-0.013	23896	3972	No	No	No	No	No	No
	(0.006)	(0.018)									
4	0.139	0.110	-0.028*	23235	3929	No	No	No	No	No	No
	(0.006)	(0.017)									
5	0.126	0.111	-0.015	23235	3929	No	No	No	No	No	No
	(0.006)	(0.017)									
6	0.112	0.087	-0.025	22657	3803	Yes	No	No	No	No	No
	(0.006)	(0.017)									
7	0.097	0.056	-0.040**	22657	3803	Yes	Yes	No	No	No	No
	(0.006)	(0.017)									
8	0.096	0.053	-0.043**	17421	3258	Yes	Yes	Yes	No	No	No
	(0.007)	(0.018)									
9	0.091	0.050	-0.040**	17421	3258	Yes	Yes	Yes	Yes	No	No
	(0.007)	(0.018)									
10	0.089	0.058	-0.031*	17421	3258	Yes	Yes	Yes	Yes	Yes	No
	(0.006)	(0.018)									
11	0.076	0.039	-0.036**	17421	3258	Yes	Yes	Yes	Yes	Yes	Yes
	(0.006)	(0.016)									

Source: Data are from SIPP, and include individuals from year 1996, 2001 and 2004. Notes: Standard errors are in parentheses. Sample includes adults' male between 22 to 65 years of age. Row 1 is the unconditional pooled OLS. Row 2 includes yearly dummies. Row 3 adds race (White, Black and Hispanic). Row 4 includes four brackets of seniority dummies, and seniority squared divide by 100 and seven brackets of years of education dummies. Row 5 adds marital status, dummies variable indicating whether the individual lives in the female head household and have children younger than 18 living in the family. Row 6 adds metropolitan, private firm and three brackets of firm size dummies. Row 7 includes dummy variable denoting possession of health insurance. Row 8 adds union dummies. Row 9 includes state dummies. Row 10 adds ten industry dummies. Row 11 includes ten occupation dummies. * indicates 90 percent statistically significant different between natives and immigrants. ** indicates 95 percent statistically significant different between natives and immigrants.

Table 3.7: Sensitivity Analysis. OLS Estimate Effect of Training on Earnings. Dependent Variable: Log of Hourly Earning

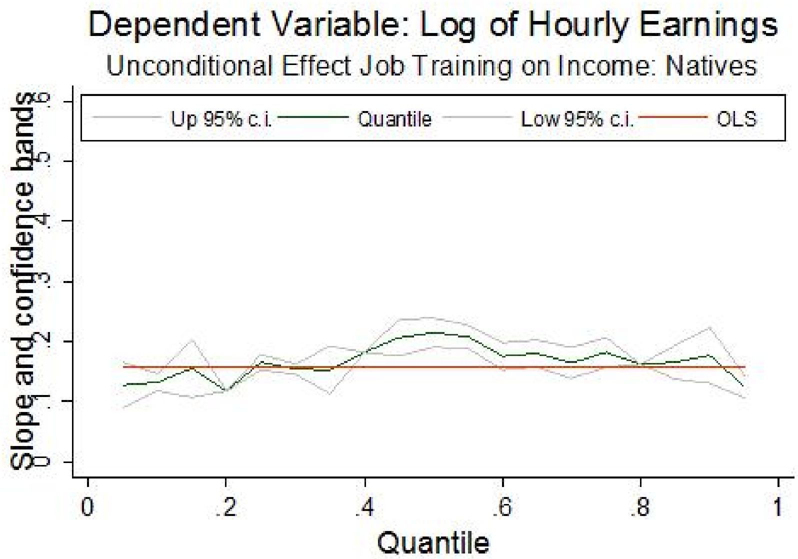
Training	Native	Im.	Diff.	Firm Size	Health Insur.	Union	State	Ind.	Occ.	N (Natives/Im.)
1	0.056	0.053	-0.003	Yes	Yes	Yes	Yes	Yes	Yes	7194 / 838
	(0.009)	(0.032)								
2	0.064	-0.134	-0.198	Yes	Yes	Yes	Yes	Yes	Yes	1599 / 233
	(0.023)	(0.073)								
3	0.080	-0.011	-0.091	Yes	Yes	Yes	Yes	Yes	Yes	7853/1140
	(0.010)	(0.029)								
4	0.068	0.070	0.003	Yes	Yes	Yes	Yes	Yes	Yes	9568/2118
	(0.008)	(0.020)								
5	0.122	0.061	-0.061	Yes	No	No	No	No	No	7853/1286
	(0.010)	(0.030)								
6	0.100	0.102	0.002	Yes	No	No	No	No	No	9568/2517
	(0.008)	(0.021)								

Source: Data are from SIPP, and include individuals from year 1996, 2001 and 2004. Sample includes adults' male between 15 to 65 years of age. Notes: Standard errors are below coefficients. All observable covariates includes yearly dummies, race (White, Black and Hispanic) dummies, seniority dummies, seniority squared divide by 100, years of education dummies, marital status dummies, dummies variable indicating whether the individual lives in the female head household, having children younger than 18 living in the family dummies, metropolitan, private firm, firm size dummies, dummy variable denoting possession of health insurance, union dummies, state dummies, industry dummies and occupation dummies. Row 1 is the pooled OLS model conditioned on all observable characteristics with sample including individuals with High School diploma. Row 2 is the pooled OLS model conditioned on all observable characteristics with sample including individuals with bachelor diploma. Row 3 is the pooled OLS model conditioned on all observable characteristics with sample including not currently married individuals. Row 4 is the pooled OLS model conditioned on all observable characteristics with sample including currently married individuals. Row 5 is the pooled OLS model conditioned on all observable characteristics except health insurance, union, state, industry and occupation dummies with sample including not currently married individuals. Row 6 is the pooled OLS model conditioned on all observable characteristics except health insurance, union, state, industry and occupation dummies with sample including currently married individuals.

Table 3.8: OLS/Matching Estimate Effect of Training on Earnings

Methodology	Native		
	Training Premium	N. Treat.	N. Control
OLS	0.076*** (0.006)		
Nearest Neighbor Matching	0.108*** (0.007)	1078	4450
Kernel Matching	0.229*** (0.005)	1078	4450
Methodology	**Immigrant**		
	Training Premium	N. Treat.	N. Control
OLS	0.039*** (0.016)		
Nearest Neighbor Matching	0.063*** (0.022)	304	2488
Kernel Matching	0.184*** (0.017)	304	2488

Note: Sample includes male workers between 15 to 65 years of age. Standard errors are in parentheses. * Statistically significant at 0.10 level; *** at the 0.05 level. For OLS and PSM, all observable covariates includes yearly dummies, race (White, Black and Hispanic) dummies, seniority dummies, seniority squared divide by 100, years of education dummies, marital status dummies, dummies variable indicating whether the individual lives in the female head household, having children younger than 18 living in the family dummies, metropolitan, private firm, firm size dummies, dummy variable denoting possession of health insurance, union dummies, state dummies, industry dummies and occupation dummies. The OLS observations for male native-born workers number 21,489 and those for male foreign-born workers 2,528.



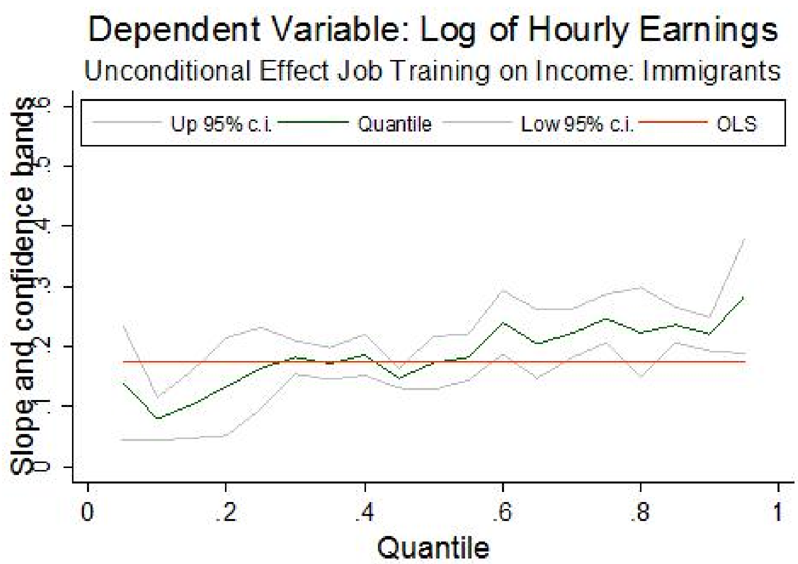


Figure 3.1: Unconditional Effect of Job Training on Earnings for Natives and Immigrants. Dependent Variable: Log of Hourly Earnings.

Source: Data are from SIPP, and include individuals from year 1996, 2001 and 2004. Sample includes adults' male between 15 to 65 years of age.

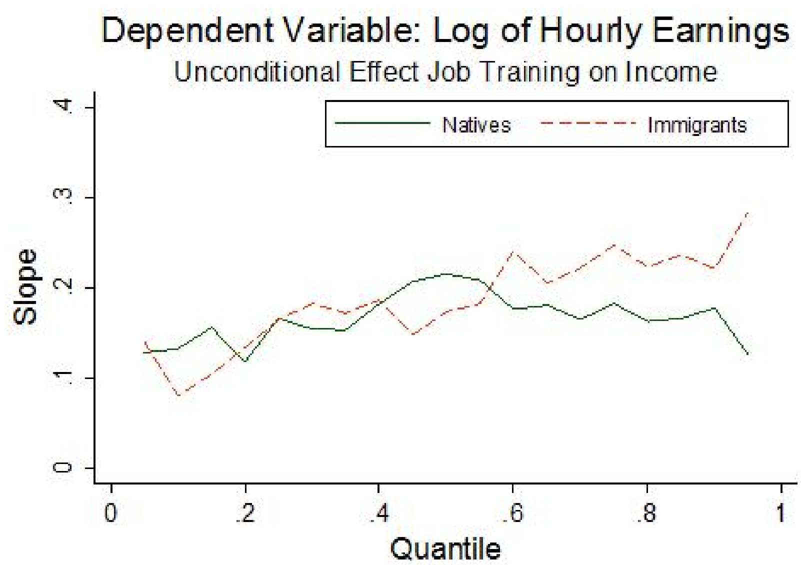


Figure 3.2: Unconditional Effect of Job Training on Earnings for Natives and Immigrants. Dependent Variable: Log of Hourly Earnings.

Source: Data are from SIPP, and include individuals from year 1996, 2001 and 2004. Sample includes adults' male between 15 to 65 years of age.

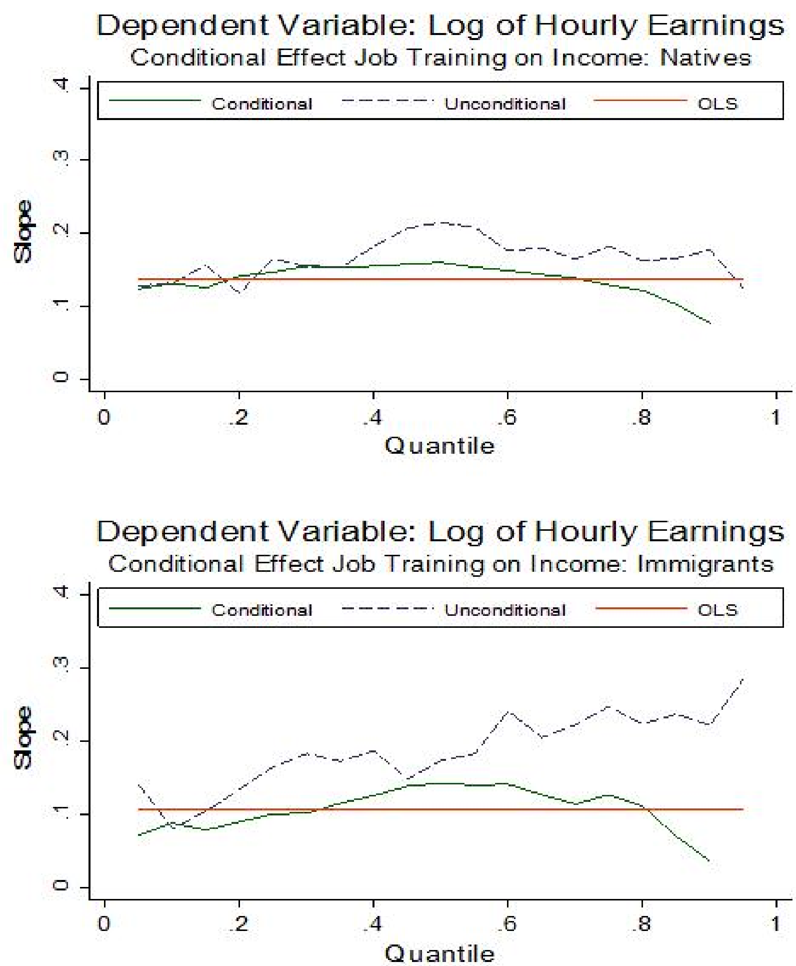
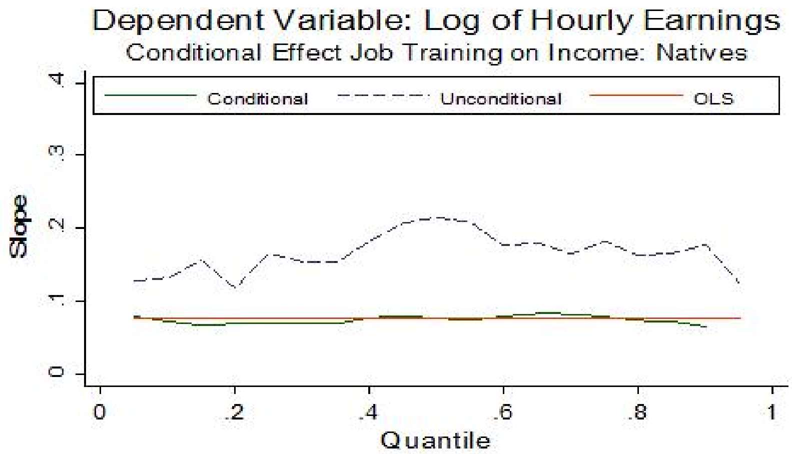


Figure 3.3: Conditional Effect of Job Training on Earnings for Natives and Immigrants (Quantiles Regression). Dependent Variable: Log of Hourly Earnings.

Source: Data are from SIPP, and include individuals from year 1996, 2001 and 2004. Sample includes adults' male between 15 to 65 years of age. Note: Quantiles regression conditioning on yearly dummies, race (White, Black and Hispanic) dummies, seniority dummies, seniority squared divide by 100, years of education dummies.



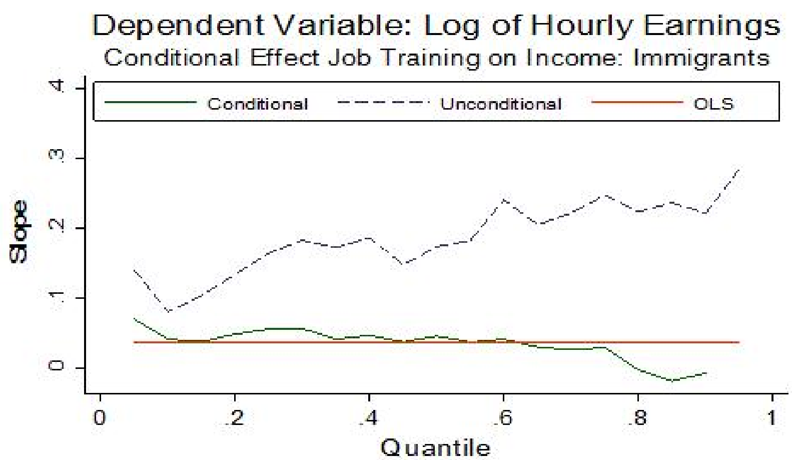
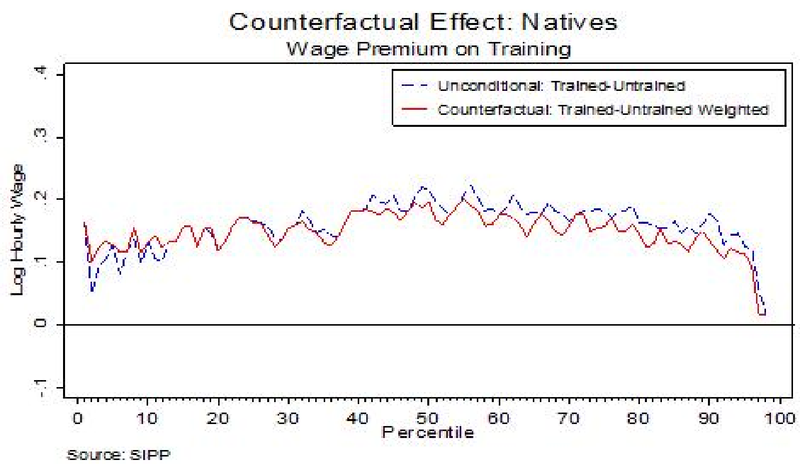


Figure 3.4: Conditional Effect of Job Training on Earnings for Natives and Immigrants (Quantiles Regression). Dependent Variable: Log of Hourly Earnings.

Source: Data are from SIPP, and include individuals from year 1996, 2001 and 2004. Sample includes adults' male between 15 to 65 years of age. Note: Quantiles regression conditioning on yearly dummies, race (White, Black and Hispanic) dummies, seniority dummies, seniority squared divide by 100, years of education dummies, marital status dummies, dummies variable indicating whether the individual lives in the female head household, having children younger than 18 living in the family dummies, metropolitan, private firm, firm size dummies, dummy variable denoting possession of health insurance, union dummies, state dummies, industry dummies and occupation dummies.



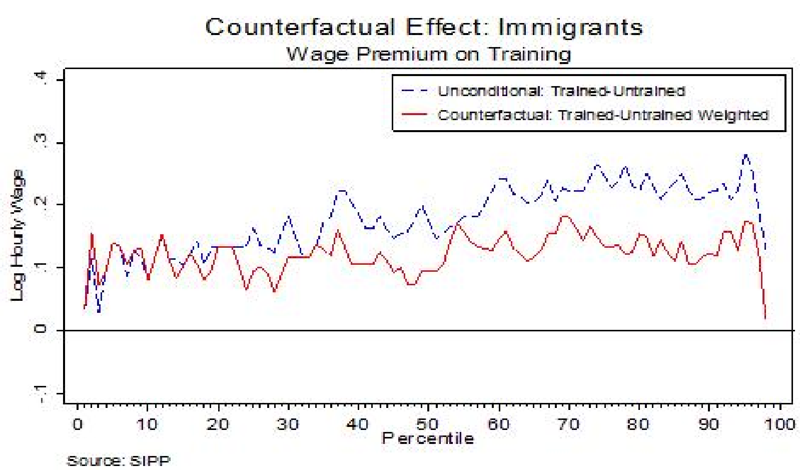


Figure 3.5: Unconditional and Counterfactual Effect of Job Training on Earnings for Natives and Immigrants (DFL Model). Dependent Variable: Log of Hourly Earnings.

Source: Data are from SIPP, and include individuals from year 1996, 2001 and 2004. Sample includes adults' male between 15 to 65 years of age. Note: DiNardo, Fortin and Lemieux (DFL) model conditions on yearly dummies, race (White, Black and Hispanic) dummies, seniority dummies, seniority squared divide by 100, years of education dummies.

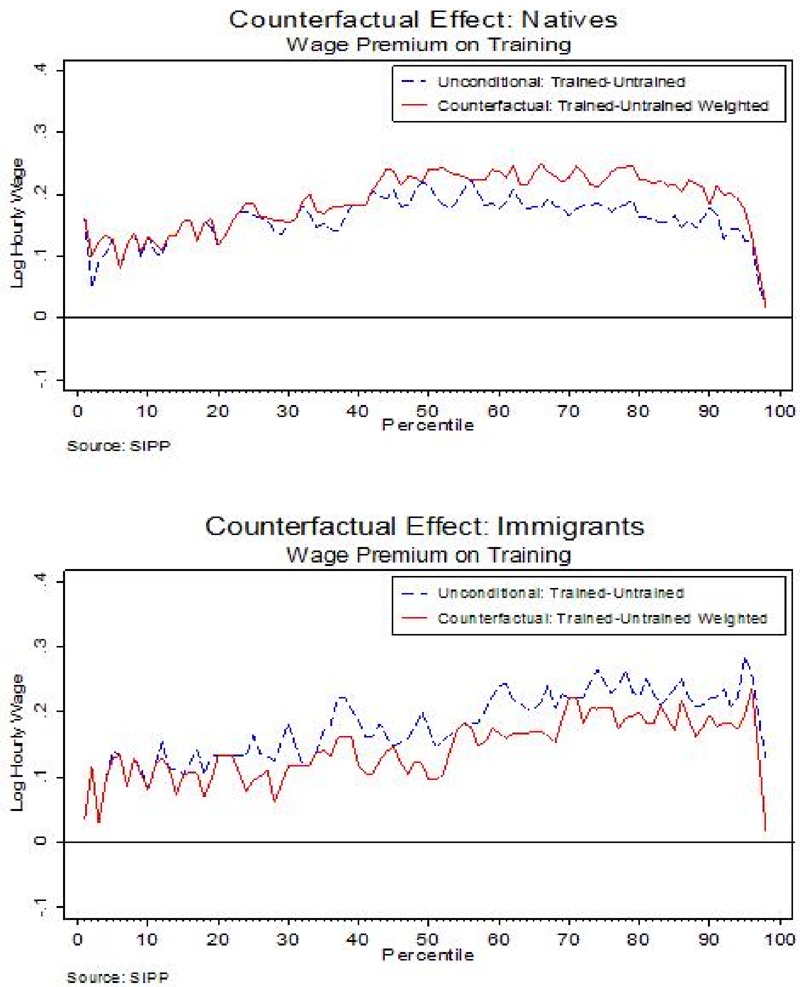


Figure 3.6: Unconditional and Counterfactual Effect of Job Training on Earnings for Natives and Immigrants (DFL Model). Dependent Variable: Log of Hourly Earnings.

Source: Data are from SIPP, and include individuals from year 1996, 2001 and 2004. Sample includes adults' male between 15 to 65 years of age. Note: DiNardo, Fortin and Lemieux (DFL) model conditions on yearly dummies, race (White, Black and Hispanic) dummies, seniority dummies, seniority squared divide by 100, years of education dummies, marital status dummies, dummies variable indicating whether the individual lives in the female head household, having children younger than 18 living in the family dummies, metropolitan, private firm, firm size dummies, dummy variable denoting possession of health insurance, union dummies, state dummies, industry dummies and occupation dummies.

Bibliography

Aakvik, A., "Bounding a Matching Estimator: The Case of a Norwegian Training Program," *Oxford Bulletin of Economics and Statistics*, 2001, *63*, 115–43.

Abadie, Alberto, Joshua Angrist, and Guido Imbens, "Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings," *Econometrica*, 2002, *70* (1), 91–117.

Ahlstrand, Amanda, Max Armbruster, Laurie Bassi, Dan McMurrer, and Mark Van Buren, "Workplace Education Investments and Strategies for Lower Wage Workers: Patterns and Practices in Employer-Provided Education," in Richard Kazis and Marc S. Miller, eds., *Low-Wage Workers in the New Economy*, The Urban Institute Press, 2001, pp. 327–45.

Aidt, T. and Z. Tzannatos, *Unions and Collective Bargaining-Economic Effects in a Global Environment*, The Word Bank,Washington D.C., 2002.

Antecol, Heather and Michael D. Steinberger, "Female Labor Supply Differences by Sexual Orientation: A Semi-Parametric Decomposition Approach," *Robert Day School of Economics and Finance Research Paper*, 2009, (2009-02).

Ashenfelter, Orley, "Determining Participation in Income-Tested Social Programs," *Journal of the American Statistical Association*, 1983, *78* (383), 517–25.

_ **and David Card**, "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs," *Review of Economics and Statistics*, 1985, *67*, 648–60.

Bassi, Laurie J., "The Effect of CETA on the Postprogram Earnings of Participants," *Journal of Human Resource*, 1983, *18* (4), 539–56.

_ **, Margaret C. Simms, Lynn C. Burbridge, and Charles L. Betsey**, *Measuring the Effect of CETA on Youth and the Economically Disadvantaged*, Washington, D.C.:Urban Institute, 1984.

Becker, Sascha O. and Andrea Ichino, "Estimation of Average Treatment Effects based on Propensity Score," *Stata Journal*, 2002, *2* (November), 358–77.

_ **and Marco Caliendo**, "Sensitivity Analysis for Average Treatment Effects," *The Stata Journal*, 2007, *7* (1), 71–83.

Blanchflower, D. G., “Changes over Time in Union Relative Wage Effects in Great Britain and United States,” *NBER Working Papers*, 1997, (6100).

_ **and R. Freeman**, “Growing into Work,” *Discussion Paper no 296*, 1996, (London School of Economics Centre for Economic Performance, London).

Bloom, Howard S., *Back to Work: Testing Reemployment Services for Displaced Workers*, W.E. Upjohn Institute for Employment Research, 1990.

Booth, A., *The Economics of the Trade Union*, Cambridge: Cambridge University Press, 1995.

Briggs, Vernon M., *Immigration and American Unionism*, Cornell University Press, 2001.

Bryant, Edward C. and Kalman Rupp, “Evaluating the Impact of Ceta On Participant Earnings,” *Evaluation Review*, 1987, *11* (4), 473–92.

Bryson, Alex, “The Union Membership Wage Premium: An Analysis Using Propensity Score Matching,” *Center for Economic Performance Discussion Paper*, 2002, *530.*

_ **, Richard Dorsett, and Susan Purdon**, “The Use of Propensity Score Matching in the Evaluation of Active Labour Market Policies,” *Working Paper for Policy Studies Institute and National Centre for Social Research*, 2002, *4.*

Budd, J. W. and I. Na, “The Union Membership Wage Premium for Employees Covered by Collective Brgaining Agreements,” *Journal of Labor Economics*, 2000, *18* (4), 783–807.

Butler, Richard and J. J. Heckman, “The Government’s Impact on the Labor Market Status of Black Americans: A Critical Review,” in Leonard J. Hausman, Orley Ashenfelter, Bayard Rustin, Richard F. Schubert, and Donald Slaiman, eds., *Equal Rights and Industrial Relations*, Industrial Relations Research Association, 1977, pp. 235–81.

Capps, Randy, Michael Fix, Jeffrey S. Passel, Jason Ost, and Dan Perez-Lopez, “A Profile of the Low-Wage Immigrant Workforce,” *Immigrant Families and Workers*, 2003, *4.*

Card, David, “The Effect of Unions on the Structure of Wages: A Longitudinal Analysis,” *Econometrica*, 1996, *64* (July), 957–79.

Chen, Vivien W. and Krissy Zeiser, *Implementing Propensity Score Matching Causal Analysis with Stata*, Population Research Institute, Penn State University, 2008.

Dabalen, A., “The Effect of Unions on Wages in South Africa: Repeated Cross-section Estimates,” *Working Paper*, 1998, (University of California, Berkeley).

Decker, Paul T. and Walter Corson, “International Trade and Worker Displacement: Evaluation of the Trade Adjustment Assistance Program,” *Industrial and Labor Relations Review*, 1995, *48* (4), 758–74.

Dehejia, Rajeev H. and Sadek Wahba, "Causal Effects in Nonexperimental Studies: Reevaluation of the Evaluation of Training Programs," *Journal of the American Statistical Association*, 1999, *94*, 1053–1062.

_ **and** _ , "Propensity Score Matching Methods for Nonexperimental Causal Studies," *Review of Economics and Statistics*, 2002, *84* (1), 151–61.

Dickinson, Katherine P., Terry R. Johnson, and Richard W. West, "An Analysis of the Impact of CETA Programs on Participants' Earnings," *Journal of Human Resources*, 1986, *21* (1), 64–91.

Donohue, John J. and Peter Siegelman, "Allocating Resources among Prisons and Social Programs in the Battle against Crime," *Journal of Legal Studies*, 1998, *27* (1), 1–43.

Edid, Maralyn, *Bridging the Gap: Training Needs Assessment of the Immigrant Workforce in Onondaga County, NY*, School of Industrial and Labor Relations, 2007.

Eren, Ozkan, "Measuring the Union-Nonunion Wage Gap Using Propensity Score Matching," *Industrial Relations*, 2007, *46* (4), 766–80.

Fairris, D., "Unions and Wage Inequality in Mexico," *Industrial and Labor Relations Review*, 2003, *56* (3), 481–97.

Fan, V. C. and J. Batalova, "Foreign-Born Wage and Salary Workers in the US Labor Force and Unions," *Migration Policy Institute*, 2007.

Farber, H. S., "Notes on the Economics of Labor Unions," *Working Paper 452*, 2001. Industrial Relations Section.

Flores-Lagunes, Alfonso, Arturo Gonzalez, and Todd Neumann, "Learning but Not Earning? The Value of Job Corps Training for Hispanic Youths," *IZA discussion paper*, 2006, *1638.*

_ **,** _ **, and** _ , "Estimating the Effects of Length of Exposure to a Training Program: The Case of Job Corps," *IZA discussion paper*, 2007, *2846.*

Fortinand, Nicole and Daniel Parent, "The Training Divide: A Canada-US Comparison of Employee Training," 2005. The paper was prepared for the joint HRSDC/Industry Canada/SSHRC Skills Research Initiative.

Foster, E. Michael, "Propensity Score Matching: An Illustrative Analysis of Dose Response," *Medical Care*, 2003, *41* (10), 1183–1192.

Frazis, H., M. Gittleman, M. Harrigan, and M. Joyce, "Results from the 1995 Survey of Employer-Provided Training," *Monthly Labor Review*, 1998, pp. 3–13.

Freeman, R. B. and J. L. Medoff, *What Do Unions Do?*, New York: Basic Books, Inc., 1984.

Frolich, Melly and Blaise Melly, "Estimation of Quantile Treatment Effects with STATA," *Working paper*, 2008.

Green, F., "The Trade Union Wage Gap in Britain:Some Recent Estimates," *Economics Letters*, 1988, *27*, 183–7.

Heckman, James, Hidehiko Ichimura, and Petra Todd, "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies*, 1997, *64* (4), 605–54.

_ , _ , and _ , "Matching as an Econometric Evaluation Estimator," *Review of Economic Studies*, 1998, *65* (2), 261–94.

_ , _ , Jeffrey Smith, and Petra Todd, "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 1998, *66* (5), 1017–1098.

_ , Robert J. LaLonde J., and Jeffrey A. Smith, "The Economics and Econometrics of Active Labor Market Programs," *Handbook of Labor Economics,*, 1999, *III* (1), 1865–2097.

Hildreth, Andrew K. G., "Union Wage Differentials for Covered Members and Nonmembers in Great Britain," *Journal of Labor Research*, 2000, *21* (winter), 133–47.

Hirsch, Barry T., "Reconsidering Union Wage Effects: Surveying New Evidence on an Old Topic," *Journal of Labor Research*, 2004, *25* (Spring), 233–66.

_ and Edward J. Schumacher, "Unions, Wages and Skills," *Journal of Human Resources*, 1998, *33* (Winter), 201–19.

k. Nakumura, H. Sato, and T. Kamiya, *Do Labor Unions really Have a Useful Role?*, Tokyo:Sago Rodo Kenkyujo, 1988.

Kane, Leonora, "Mentoring for Black Students: Improving Equality of Opportunity," *Education + Training*, 1994, *36* (8), 18–24.

Koenker, Roger and Jr Gilbert Bassett, "Regression Quantiles," *Econometrica*, 1978, *46* (1), 33–50.

Lalonde, Robert, "Evaluating the Econometric Evaluations of Training Programs," *American Economic Review*, 1986, *76* (4), 604–20.

Lewis, H. G., "Uion Relative Wage Effects," *Handbook of Labor Economics*, 1986, *II*, 1139–1181.

Main, B. G. M. and B. Reilly, "Women and the Union Wage Gap," *The Economic Journal*, 1992, *102* (410), 49–66.

Mantel, N. and W. Haenszel, "Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease," *Journal of the National Cancer Institute*, 1959, *22*, 719–48.

Milkman, R., "Labor Organizing among Mexicon-born Workers in the United States-Recent Trends and Future Prospects," *Labor Studies Journal*, 2007, *32* (1), 96–112.

Moll, P. G., "Black South African Unions: Relative Wage Effects in International Perspective," *Industrial and Labor Relations Reviews*, 1993, *46* (2), 245–61.

Orr, L. L., H. S. Bloom, S.H. Bell, F. Doolittle, W. Lin, and G. Cave, *Does Training for the Disadvantaged Work?*, The Urban Institute, 1996.

Pagan, Adrian and Aman Ullah, *Nonparametric Econometrics*, Cambridge University Press, 1999.

Panagides, A. and H. A. Patrinos, "Union-Nonunion Wage Differentials in the Developing World, A Case Study of Mexico," *Policy Research Working Paper*, 1994, *1269.*

Parent, Daniel, "Labour Market Outcomes and Schooling in Canada: Has the Value of a High School Degree Changed over Time," *CIRANO Working Papers*, 1999, *99* (42).

Park, Peter Y. and Frank F. Saccomanno, "Reducing Treatment Selection Bias for Estimating Treatment Effects Using Propensity Score Method," *Journal of Transportation Engineering*, 2007, *133* (2), 112–18.

Patrinos, H. A. and C. Sakellariou, "North American Indians in the Canadian Labour Market: a Decomposition of Wage Differentials," *Economics of Education Review*, 1992, *11* (3), 257–66.

Rosenbaum, Paul R., *Observational Studies*, New York: Springer, 2002.

_ **and Donald B. Rubin**, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 1983, *70*, 41–55.

Schochet, Peter Z., John Burghardt, and Steven Glazerman, "The Impacts of Job Corps on Participants' Employment and Related Outcomes," Technical Report 8140-530, Mathematica Policy Research 2001.

Simpson, W., "The Impact of Unions on the Structure of Canadian Wages: An Empirical Analysis with Microdata," *Canadian Journal of Economics*, 1985, *18* (1), 164–81.

Smith, James P. and Finis Welch, *Closing the Gap: Forty Years of Economic Progress for Blacks*, The RAND Corporation, 1986.

Solivas, Emeterio S., Girly M. Ramirez, and Allan N. Manalo, "The Propensity Score Matching for Correcting Sample Selection Bias," *10th National Convention on Statistics (NCS)*, 2007.

Standing, G., "Do Unions Impede or Accelerate Structural Adjustment? Industrial versus Company Union in an Industrialising Labour Market," *Cambridge Journal of Economics*, 1992, *16*, 327–54.

Teal, F., "The Size and Source of Economic Rents in a Developing Country Manufacturing Labour Market," *The Economic Journal*, 1996, (106), 963–76.

Tumlin, Karen C. and Wendy Zimmermann, "Immigrants and TANF: A Look at Immigrant Welfare Recipients in Three Cities," *Assessing the New Federalism Occasional Paper*, 2003, *69.*

Waldinger, R. and C. Der-Martirosian, "Immigrant Workers and American Labor: Challenge¡ or Disaster," in "Organizing Immigrants-The Challenge for Unions In Contemporary California," Cornell University Press, 2000.